République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université 8 Mai 1945 Guelma



Faculté des Sciences et de la Technologie
Département Electronique et Télécommunications
LABORATOIRE PROBLEMES INVERSES,
MODELISATION, INFORMATION ET SYSTEMES (PIMIS)

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT EN 3ème CYCLE

Domaine : Sciences et Technologie. **Filière** : Electronique
Spécialité : Signaux et Images Biométriques

Présentée par

Khider Ali

Intitulée

Un système biométrique multimodal basé sur la fusion visage-iris

Soutenue le : 02 Mars 2023

Devant le Jury composé de :

Mr Bourouba Hocine	Professeur	Univ. de Guelma	Président
Mr Djemili Rafik	Professeur	Univ. de Skikda	Encadreur
Mr Messikh Lotfi	Professeur	Univ. de Skikda	Examinateur
Mr Doghmane Hakim	MCA	Univ. de Guelma	Examinateur
Mme Boucerredj Leila	MCA	Univ. de Guelma	Examinateur

Année Universitaire **: 2022-2023**

الجمهوريـة الجزائريـة الديمقراطيـة الشعبيـة

وزارة التعليم العالي والبحث العلمي

جامعة 8 ماي 1945 قالمـة



كلية العلوم والتكنولوجيا

قسم الالكترونيك والاتصالات السلكية واللاسلكية

مخبر المشاكل العكسية والنمذجة والمعلومات والأنظمة (PIMIS)

أطروحة

لنيل شهادة الدكتوراه في الطور الثالث

الميدان: علوم وتكنولوجيا. **الشعبة**: الكترونيك

الاختصاص: إشارات وصور بيومترية

من إعداد:

علي خيذر

بعنوان

نظام بيومتري متعدد الوسائط يعتمد على الدمج بين الوجه والقزحية

بتاريخ: 02 مارس 2023

أمام لجنة المناقشة المكونة من:

السيد بوروبة حوسين	أستاذ جامعي	بجامعة قالمة	رئيسا
السيد جميلي رفيق	أستاذ جامعي	بجامعة سكيكدة	مشرفا
السيد مسيخ لطفي	أستاذ جامعي	بجامعة سكيكدة	ممتحنا
السيد دغمان حكيم	أستاذ محاضر (أ)	بجامعة قالمة	ممتحنا
السيدة بوسراج ليلى	استاذة محاضرة (أ)	بجامعة قالمة	ممتحنا

السنة الجامعية: 2022–2023

Dédicaces



À Tous …

À ma famille

Ma mère et Mon père

Mon frère et Mes sœurs

Ma Femme et Ma fille

Pour leur gentillesse, leur disponibilité, leur encouragement…

Qu'il m'a prodiguée.

À mes amis

Pour leurs conseils, leur aide,… aussi pour leur amitié

Remerciements

Je remercie en premier lieu mon grand DIEU qui m'a donné à la fois le courage, la volonté, et la patience afin d'élaborer cette thèse de recherche scientifique.

Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.

Je tiens à remercier Monsieur **Rafik Djemili**, Professeur à l'Université de Skikda et membre dans notre laboratoire PIMIS qui m'a encadré tout au long de cette thèse, et qui m'a fait partager ses brillantes intuitions.

Je tiens à remercier Monsieur **Abdelhani Boukrouche** رحمه الله وأدخله فسيح جنانه, Professeur à l'Université 08 Mai 1945 de Guelma, et l'ancien directeur de notre laboratoire PI:MIS pour sa gentillesse, sa disponibilité permanente et pour les nombreux encouragements qu'il m'a prodiguée.

J'exprime ma profonde gratitude à Monsieur. **Bourouba Hocine**, Professeur à l'Université 08 Mai 1945 de Guelma, et le directeur de notre laboratoire PIMIS pour son aide sur le plan scientifique que pour sa sympathie et sa disponibilité.

Je remercie vivement Monsieur **Bourouba Hocine**, Professeur à l'Université de Guelma, pour l'honneur qu'il me fait, en acceptant la présidence de ce jury.

J'adresse mes sincères remerciements à Monsieur **Messikh Lotfi**, Professeur à Skikda, Monsieur **Doghmane Hakim**, MCA à l'Université de Guelma et Madame **Boucerredj Leila**, MCA à l'Université de Guelma, pour avoir accepté de participer à mon jury de thèse et l'honneur qu'ils m'ont fait en acceptant d'examiner ce travail.

Je tiens enfin à remercier l'ensemble des membres du laboratoire PIMIS : chercheurs, personnels administratifs et techniques, et doctorants, pour leurs conseils, leur aide technique mais aussi pour leur amitié

Table des matières

Liste des tableaux

Liste des figures

Liste des abréviations

2D-DWT…………..	2D Discrete Wavelet Transform
CLAHE……………	Contrast Limited Adaptive Histogram Equalisation
GMPTR…………...	Gabor Magnitude-Phase-Based Texture Representation
NLDA……………...	Null-space Linear Discriminant Analysis
GLL……………......	Gabor LBP LPQ
TWSBF……………	Top-level's Wavelet Sub-Bands
PCA……………......	Principal Component Analysis
LDA……………......	Linear Discriminant Analysis
STFT……………....	Short-Term Fourier Transform
LBPH……………...	Local Binary Pattern Histogram
SVM…………….....	Support Vector Machine
LBP……………......	Local Binary Pattern
RLP……………......	Robust Linear Programming
LPQ……………......	Local Phase Quantization
ORL……………......	Olivetti Research Laboratory
PIMIS……………....	Laboratory of Inverse Problems, Modeling, Information and Systems
RGB……………......	The color space (Red, Green, Blue)
PIN…………….......	Personal Identification Number
ROI…………….......	Region Of Interest
SVM……………......	Support Vector Machines
KNN……………......	k-Nearest Neighbors
Part……………......	Partitionnement
Chv……………......	Chevauchement
Norm…………….....	Normalisation
GL…………….........	Gabor-LDA
BL………………....	LBP-LDA
QL……………........	LPQ-LDA

INTRODUCTION GENERALE

Au cours des dernières décennies, l'utilisation croissante de nombreuses techniques développées de piratage et de falsification conduit à une demande croissante des méthodes de reconnaissance de l'identité d'une personne [1]. Les systèmes biométriques sont des technologies en constante évolution et prometteuses qui peuvent être utilisées dans des systèmes automatiques pour identifier et/ou vérifier l'identité d'une personne de manière unique et efficace sans que l'utilisateur ait besoin de porter ou de se souvenir de rien, contrairement aux méthodes traditionnelles. Ces systèmes ont été largement utilisés dans nombreuses applications sensibles gouvernementales et civiles, en particulier celles comprenant un contrôle d'accès automatique à des lieux physiques ou virtuels, tels que les points de passage des frontières, les machines ATM, les systèmes de sécurité et de surveillance, les transactions bancaires, la sécurité des ordinateurs/réseaux, …etc. [2].

L'identification personnelle basée sur des caractéristiques biométriques a de nombreux avantages par rapport aux méthodes traditionnelles basées sur la mémoire (par exemple, les mots de passe ou les numéros d'identification personnels (PIN)) ou les méthodes basées sur des pièces physique (par exemple, le permis de conduire, le passeport, la carte d'identité ou une simple clé), car elles sont difficiles à transférer, à perdre, à oublier ou à dupliquer. En plus, l'utilisation de la biométrie dans la tâche d'identification d'une personne est plus pratique et plus agréable à utiliser que les méthodes traditionnelles, de sorte que les utilisateurs n'ont pas besoin de se souvenir ou de transporter quoi que ce soit avec eux, évitant ainsi la perte, le partage ou l'oubli d'informations personnelles. Enfin, le niveau de sécurité atteint grâce aux systèmes biométriques peut être plus élevé que celui obtenu par les méthodes traditionnelles [3].

De manière générale, les systèmes biométriques peuvent être divisés en deux grands types : les systèmes biométriques uni-modaux et multimodaux. Les systèmes biométriques uni-modaux sont basés sur l'utilisation d'une seule source d'information, par exemple l'empreinte digitale, la géométrie de la main, l'iris, la rétine, le visage, l'empreinte palmaire, la géométrie de l'oreille, l'ADN, la voix, la démarche, la signature ou encore la dynamique de frappe au clavier, sont autant de modalités biométriques différentes pour établir l'identité d'une personne. Par contre, les systèmes biométriques multimodaux utilisent plus d'une modalité [3], [4].

L'emploi de plus d'une modalité, échantillon, capteur ou algorithme biométrique pour obtenir une reconnaissance, communément appelée multi-biométrie, est une technique qui gagne rapidement en popularité. En incorporant la multi-biométrie dans le processus de

reconnaissance, de nombreux défauts des systèmes traditionnels à uni-biométrie peuvent être atténués et la précision globale de la reconnaissance peut être améliorée. La multi-biométrie peut augmenter la robustesse du système en supprimant la dépendance à une approche biométrique particulière. De plus, un système qui utilise plus d'une caractéristique biométrique peut être plus difficile à falsifier [3]. Deux modalités biométriques deviennent de plus en plus populaires parmi la communauté des chercheurs en raison de leur acceptabilité, de leur faible coût de déploiement, de leur fiabilité, mais aussi et surtout la connexion naturelle entre elles, qui sont la reconnaissance par le visage et l'iris, offrant par ailleurs la possibilité d'utiliser un seul capteur. Pour cela, nous avons proposé un système multimodal qui utilise le visage et l'iris, multi-algorithme qui combine des méthodes de reconnaissance tel que la méthode de l'égalisation d'histogramme adaptative limitée par le contraste (CLAHE), les filtres de Gabor, Motif binaire local (LBP), Quantification par phase locale (LPQ), L'analyse discriminante linéaire (LDA), et l'apprentissage approfondi (Autoencoder). Le système biométrique proposé, nous a fourni des résultats fiables et précis, outre qu'il comprend de nombreuses caractéristiques, en plus, il offre plus de sécurité que les systèmes n'utilisant qu'une seule modalité.

Buts et objectifs de la thèse

L'objectif principal de cette thèse de recherche est la conception et la mise en œuvre d'un système d'identification biométrique multimodal et hybride (multi-biométrique, multi-instances, et multi-algorithme), permettant de trouver l'identité d'une personne à l'aide d'une combinaison de traits biométriques du visage et de l'iris. Ces deux modalités biométriques se complètent, au sens que le visage est le moins intrusif et l'iris est le plus précis, et qui sont naturellement connectés donnant la possibilité d'utiliser le même capteur à distance. Les objectifs spécifiques sont la conception et l'adaptation des algorithmes biométriques unimodaux, robustes et prêts à l'emploi pour un système multimodal qui intègre les points forts de ces deux modalités biométriques. D'autres objectifs sont d'augmenter le degré de flexibilité du système multimodal proposé et d'éliminer la limitation imposée par le trait manquant en sélectionnant un schéma de fusion approprié. Afin d'atteindre les principaux objectifs de la recherche, les tâches suivantes doivent être étudiées :

1. Examiner et analyser l'état de l'art des approches de la reconnaissance biométrique pour les deux modalités concernés (le visage et l'iris). Ces études comprennent plusieurs étapes

du système biométrique liées au prétraitement, à l'extraction et la réduction des caractéristiques, à la méthodologie de fusion et à la classification. Dans cette tâche, l'accent sera mis sur la théorie fondamentale, les inconvénients et les hypothèses formulées par les travaux précédents liées avec notre recherche.

2. Etudier et proposer des améliorations pour développer un système multimodal puissant qui peut répondre aux exigences des applications du monde réel. Des algorithmes de reconnaissance de visage entièrement automatisé et robuste sont proposé dans lequel une représentation faciale compacte et discriminative est apprise à partir d'images de visage, qui sont prises dans des conditions totalement sans contraintes. Dans cette thèse, un certain nombre des problèmes de reconnaissance faciale les plus difficiles sont abordés, notamment le manque d'échantillons, les changements d'éclairage, d'expressions, de l'âge, l'occlusion et les différentes poses.

3. La conception d'un prototype pratique et d'un système de reconnaissance de l'iris robuste. Cette tâche implique plusieurs problèmes : (i) Proposition d'une méthode efficace pour la localisation de l'iris et la séparation de sa région de l'arrière-plan. (ii) Proposition d'une nouvelle méthode d'extraction de caractéristiques discriminantes de la région localisée de l'iris sans aucune connaissance sur la position de l'iris dans d'image, qui diffèrent des travaux précédents qui dépendent de caractéristiques extraites [5]–[9]. Cette idée est motivée par les résultats intéressants obtenus en appliquant les approches d'apprentissage approfondi pour obtenir des caractéristiques compactes et discriminantes qui sont utilisés fortement dernièrement dans la reconnaissance du visage. (iii) Proposition d'une nouvelle approche de comparaison, qui pourrait améliorer la précision des approches traditionnelles de comparaison binaire de l'iris, telles que la distance de Hamming.

4. Conception et l'implémentation d'un système multi-instance robuste basé sur la biométrie de l'iris de l'œil droit et de l'œil gauche de la même personne.

5. Investigations sur la combinaison les deux modalités (visage-iris) dans un système d'identification personnelle multi-biométrique. Cette tâche consiste à étudier des approches de fusion efficaces pour les caractéristiques biométriques adoptées.

6. Réaliser des expériences approfondies pour évaluer l'efficacité des approches et des méthodologies proposées, dans différents scénarios, sur des bases de données biométriques unimodal et multimodal à grande échelle et sans contraintes. En de plus, nous avons évalué

supériorité des systèmes unimodaux et multimodaux proposés par rapport aux systèmes de l'état de l'art.

Contributions de la thèse

Dans cette section, les principales contributions de cette recherche sont soulignées. La première partie de cette thèse aborde les problèmes des systèmes biométriques unimodaux basés sur le visage et l'iris, tandis que la seconde partie traite du problème de la conception et de la mise en œuvre d'un système d'identification biométrique multimodal hybride, permettant d'identifier d'une personne en utilisant une combinaison des traits biométriques du visage et de l'iris. Les principales contributions de cette thèse sont les suivantes :

1. Cette thèse propose une nouvelle méthode de reconnaissance des formes utilisant une combinaison de caractéristiques et l'apprentissage approfondi. Au lieu d'utiliser les images de visages dans leur espace d'origine (RGB ou Gray) comme entrée de l'apprentissage approfondi, nous proposons une combinaison de plusieurs caractéristiques extraites à partir les images d'origine, afin de créer un nouvel espace de représentation pour l'entraînement de l'apprentissage approfondi.

2. Proposée une approche multi-algorithme pour l’extraction des caractéristiques combinant des méthodes locales et globales (hybride) utilisant les filtres de Gabor, LBP et LPQ en parallèle pour assurer une représentation plus discriminante et pour que le système soit généraliste (fonctionnant avec les autre modalités), et invariant dans des environnements non contrôlés.

3. Emploi de l’algorithme de l’égalisation d'histogramme adaptative limitée par le contraste (CLAHE) pour le prétraitement.

4. Utilisation d’une seule image acquise par le même capteur à distance au fin de la reconnaissance multimodale des personnes par le visage et l’iris (application sur CASIA-IrisV4-Distance)

Plan de la thèse

Cette thèse est organisée en quatre chapitres. Un aperçu de ces chapitres est présenté comme suit.

Chapitre 1 présente les fondements des systèmes biométriques, les principales caractéristiques des modalités biométriques, les limites des systèmes biométriques unimodaux, et les avantages des systèmes multimodaux.

Chapitre 2 présente une revue des travaux antérieurs relatifs à notre recherche dans la reconnaissance des visages. Traite du problème de la reconnaissance des visages dans l'environnement non contrôlé dans la tâche d'identification, et présente un nouveau système multi-algorithme de reconnaissance des visages. En particulier, tout le système incluant les méthodes de prétraitement, d'extraction de caractéristiques, de réduction, et de classification des visages basé sur la construction de la module d'apprentissage approfondi (Autoencoder) sont présentées.

Chapitre 3 présente une revue des travaux antérieurs relatifs à la reconnaissance de l'iris, étudie le problème de la reconnaissance de l'iris dans une tâche d'identification, présente un système de reconnaissance de l'iris efficace basé sur l'extraction de la texture de l'iris, et la classification par le module proposé de l'apprentissage approfondi (Autoencoder). La deuxième partie de ce chapitre présente le schéma général multimodal hybride proposé, basé sur la fusion des scores de correspondance générés à partir des traits biométriques du visage et de l'iris au niveau du score. Dans ce chapitre, différents types de systèmes biométriques multimodaux sont également proposés, sur la base des critères d'identification de l'utilisateur.

Chapitre 4 présente les études expérimentales, et les bases de données utilisées. Fournit les résultats de la technique proposée sur le visage, l'iris, et aussi sur la fusion multimodale visage-iris. Des comparaisons avec des récents et similaire recherches sont également présentés dans ce chapitre.

Cette thèse est terminée par une conclusion générale où elle exposée également les perspectives à ce travail de recherche.

Chapitre I

Introduction à la Biométrie

I.1 Introduction

La biométrie est la science qui permet de reconnaître l'identité d'une personne sur la base de ses caractéristiques physiologiques, chimiques ou comportementales, telles que: le visage, l'iris, l'odeur, la façon de marcher ou la signature électronique…etc. Avec la nécessité des techniques robustes de reconnaissance humaine dans les applications critiques, telles que: le contrôle d'accès sécurisé, le passage des frontières internationales et les applications légales, la biométrie se positionne comme une technologie viable qui peut être intégrée dans les systèmes de management d'identité à grande échelle [10].

Dans ce chapitre, nous définissons le terme Biométrie, ses caractéristiques, modalités ainsi que ses domaines d'application.

I.2 La biométrie : notions de bases et généralités

Le mot biométrie désigne dans un sens très large l'étude quantitative des êtres vivants, mais dans notre contexte plus précis de reconnaissance et d'identification d'individus, il existe deux définitions principales qui se complètent [11]:

1. La biométrie est la science qui étudie à l'aide des mathématiques, les variations biologiques à l'intérieur d'un groupe déterminé
2. Toute caractéristique physique ou trait personnel automatiquement mesurable, robuste et distinctif qui peut être employé pour identifier un individu ou pour vérifier l'identité qu'un individu affirme

I.2.1 Caractéristique de la biométrie

Une caractéristique biométrique est une donnée contenant l'essentiel d'informations permettant de différencier deux individus. Pratiquement n'importe quelle caractéristique physiologique ou comportementale peut être considérée comme une caractéristique biométrique appelée aussi modalité, dans la mesure où elle est [12] :

– **universelle** : existe chez tous les individus,
– **unique** : permet de différencier un individu par rapport à un autre,
– **permanente** : stable dans le temps,
– **enregistrable :** collecte les caractéristiques d'un individu avec son accord,
– **mesurable :** autoriser une comparaison future.

I.2.2 Les modalités biométriques

Chaque caractéristique biométrique qui satisfait aux conditions précédentes, peut être utilisée pour identifier un individu, elle est appelé modalité biométrique, et se classer selon le type : physiologique ou comportementale, comme le montre la Figure I.1.

- **Les modalités physiologiques ou morphologiques:** Ces modalités sont uniques et permanentes, leur principe est basée sur l'identification de traits physiques particuliers d'une personne, tels que : Le visage, L'iris, L'oreille, La rétine, L'ADN, La géométrie de la main, L'empreinte palmaire, L'empreinte digitale. Ces éléments ont l'avantage d'être stables dans la vie d'un individu
- **Les modalités comportementales :** Ces modalités se basent sur l'analyse de certains comportements d'une personne comme : La démarche, La voix, La signature, La dynamique de frappe au clavier.

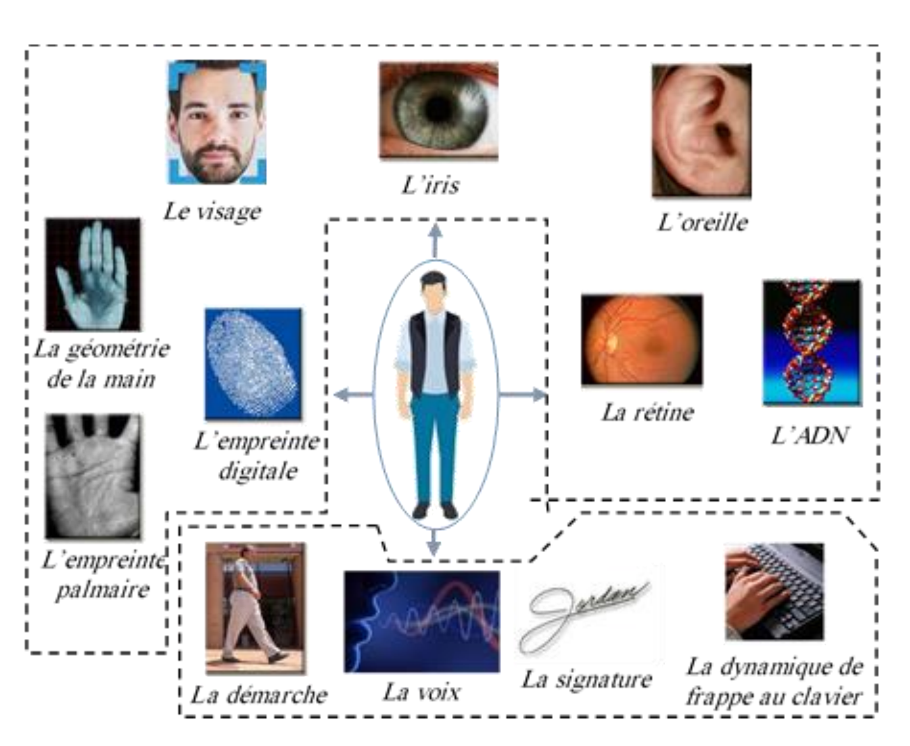


Figure I.1 Exemples de modalités biométriques (physiologiques et comportementales)

I.2.3 Comparaison entre les différentes modalités biométriques

Le Tableau I.1 compte les avantages et les inconvénients des différentes modalités biométriques. Alors que le Tableau I.2 présente une comparaison des modalités biométriques existantes en fonction de ces caractéristiques [13]:

Tableau I.1 Avantages & inconvénients des différentes modalités

Modalités	**Avantages**	**Inconvénients**
Iris	✓ L'iris recèle plus de données que les empreintes digitales ✓ Stable durant toute la vie d'une personne. ✓ Technique fiable	✗ Les effets d'illumination et d'occlusion (les yeux bougent quand on capture l'image). ✗ La qualité de l'image n'est pas bonne. ✗ Les contours de la pupille et de l'iris ne sont pas circulaires ✗ Nécessite des dispositifs de détection spéciaux.
Visage	✓ Technique moins cher. ✓ Visage capturé à distance ✓ Technique simple	✗ Technologie sensible à : • Les variations d'âge • À l'utilisation des artifices (moustaches, barbe, lunettes…) • Les variations expression et poses • variations d'illumination
Empreintes digitales	✓ Laissons derrière nous à chaque fois que nous touchons un objet. ✓ Nous aider à la recherche sur une scène de crime ✓ Elles sont fiables et inchangeable durant la vie d'un individu.	✗ Difficulté de lire l'empreinte digitale pour les travailleurs manuels. ✗ Images à faible contraste ✗ Mauvaise acquisition d'image ✗ Nécessite dispositifs de détection spéciaux. ✗ Nécessitant un contact physique
Signature	✓ Plus confortable ✓ Utilisé dans le document administratif. ✓ Accepter par les personnes ✓ Rapide et efficace	✗ Besoin d'une tablette graphique ✗ On ne peut pas utiliser à contrôle d'accès extérieur ✗ Les signatures falsifiées (imitation)
Voix	✓ Efficace au téléphone ✓ Rapide et efficace ✓ acceptées par les personnes	✗ Sensible au bruit ambiant. ✗ Voix enregistrées ✗ Sensible à l'état physique et émotionnel de l'individu

Tableau I.2 Comparaison des modalités biométriques (H=Haut, B=Bas et M=Moyenne)

Modalités	Universalité	Unicité	Permanence	Performance	Acceptabilité	Contre la falsification
L'iris	H	H	H	H	B	H
Le visage	H	B	M	B	H	B
L'ADN	H	H	H	H	B	B
La voix	M	B	B	B	H	B
La signature	L	B	B	B	H	B
l'oreille	M	M	H	M	H	M
La rétine	H	H	M	H	B	H
La démarche	M	B	B	B	H	M
L'empreinte digitale	M	H	H	H	M	H

I.2.4 Choix d'une modalité biométrique

Le choix d'une modalité dépend de sa nature, de son niveau de sécurité, ainsi que de l'environnement de leur usage (facilité d'emploi, d'analyse, de stockage, et de vérification) Le groupe international biométrique (IBG) [http://www.biometricgroup.com] offre une comparaison de différentes modalités basées sur 4 critères selon l'analyse de Zephyr (Figure I.2)

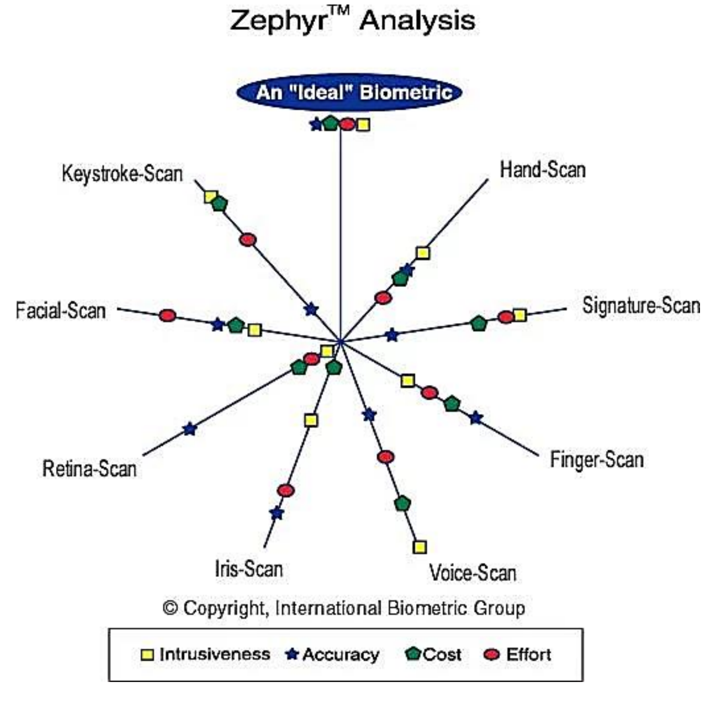


Figure I.2 Analyse de Zephyr : Critères de choix des modalités biométriques

- **Effort** : effort requis par l'utilisateur lors de l'authentification ou l'identification.
- **Coût (Cost)**: lié aux dispositifs technologiques (lecteurs, capteurs, etc…).
- **Exactitude (Accuracy)** : efficacité de la méthode (liée au taux de reconnaissance).
- **Intrusion (Intrusiveness)** : décrit l'information sur l'acceptation du système par les utilisateurs.

I.3 Le système de la reconnaissance biométrique

Un système biométrique est essentiellement un système de classification et de reconnaissance qui capture un trait biométrique d'un individu, extrait un ensemble de caractéristiques discriminantes du trait capturé, compare l'ensemble de caractéristiques extraites à un ensemble (ou à des ensembles) de modèles stockés dans la base de données du système. Ensuite, la décision finale est prise en fonction des résultats de cette comparaison. Comme le montre la Figure I.3, la structure globale de tout système biométrique se compose de quatre étapes principales, qui fonctionnent de manière séquentielle pour obtenir le résultat du système [4], [14].

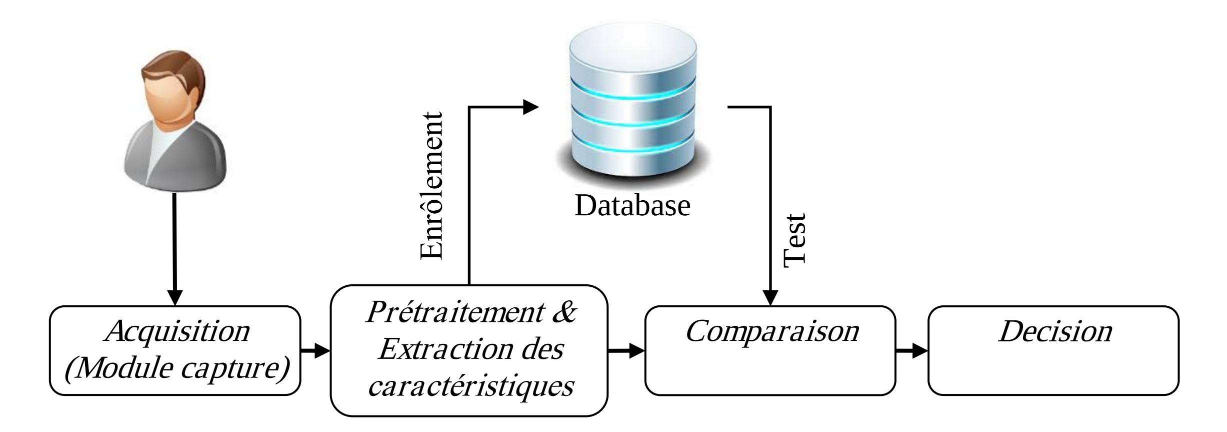


Figure I.3 La structure globale d'un système biométrique

I.3.1 Module de capture

Ce module est responsable de l'acquisition des traits biométriques d'un individu cela peut être un appareil photo, un lecteur d'empreintes digitales, une caméra de sécurité. Ainsi, dans certains systèmes biométriques, cette étape est associée à une procédure de contrôle qualité.

I.3.2 Module de prétraitement et d'extraction de caractéristiques

Typiquement, le trait biométrique capturé est soumis à des opérations de prétraitement afin d'atteindre une qualité requise, ces opérations peuvent être divisées en trois étapes. Une étape d'évaluation pour évaluer la qualité du trait biométrique capturé, sur la base d'un seuil prédéfini : soit utiliser le trait pour le traitement suivant, soit tenter de recapturer le trait à nouveau. Puis une étape de segmentation, dans laquelle la région d'intérêt du trait biométrique est séparée du bruit de fond, par exemple en détectant la région de l'iris dans l'image de l'œil. Enfin, un certain nombre d'algorithmes d'amélioration peuvent être appliqués sur la région biométrique détectée (par exemple, les algorithmes de transformation, l'égalisation d'histogramme, etc.) pour améliorer sa qualité en réduisant le niveau de bruit introduit par la caméra et les variations d'éclairage.

Après que certaines opérations de prétraitement ont été appliquées, un ou plusieurs algorithmes doivent être appliqués pour extraire seulement l'information pertinente afin de former une nouvelle représentation des données discriminante, appeler vecteur de caractéristiques, par exemple : L'Analyse des Composantes Principales (ACP), Local Binary Patterns (LBP), etc ..

Ensuite, l'ensemble des caractéristiques extraites est soit envoyé au module de comparaison pour l'identification de l'utilisateur, ou soit stocké dans la base de données du système en tant qu'un modèle lors d'une phase d'enrôlement. Idéalement, les caractéristiques extraites devraient avoir des variations minimales intra-classes et maximales inter-classes.

I.3.3 Module de comparaison

Ici, le vecteur de caractéristiques qui fournit par le module d'extraction de caractéristiques est comparé à tous les modèles précédemment stockés dans la base de données du système à travers la phase de test pour générer des scores de correspondance. Ces scores sont des mesures de similarité, auquel cas un score (distance) plus élevé indique une probabilité de dissimilitude plus élevé, un score inférieur indique une correspondance (similarité) plus étroite entre l'utilisateur et le modèle comparé.

I.3.4 Module de décision

Il s'agit de la dernière étape du système biométrique, dans laquelle, l'utilisateur est identifié ou une identité revendiquée est soit acceptée (autorisant l'utilisateur), soit rejetée (n'autorisant pas l'utilisateur) sur la base du score généré par le module de comparaison.

I.3.5 Modes de fonctionnement d’un système biométrique

Les systèmes biométriques peuvent fournir trois modes de fonctionnement, à savoir, l’enrôlement, l’authentification (ou vérification) et l’identification.

Mode enrôlement : C’est la première phase de tout système biométrique, il s’agit de l’étape pendant laquelle les représentations numériques des caractéristiques extraites du trait biométrique capturé de la personne à l'aide d'un capteur sont enregistrés dans la base de données du système, avec certaines informations biographiques, par exemple, nom, code PIN, N°ID, … etc., permettant de distinguer l'utilisateur. Ces modèles sont conservés chiffrés dans certaines applications sensibles pour des raisons de sécurité et de confidentialité [15].

Mode vérification : En mode vérification, le système valide l'identité d'une personne en comparant les données biométriques capturées avec ses propres modèles biométriques stockés dans la base de données du système. Dans un tel système, une personne qui souhaite faire reconnaître une identité, généralement via un numéro d'identification personnel (PIN), un nom d'utilisateur ou une carte à puce, et le système effectue une comparaison individuelle pour déterminer si la demande est vrai ou non (par exemple, Ces données biométriques appartiennent-elles à Salah ?). La vérification d'identité est généralement utilisée pour la reconnaissance positive, où le but est d'empêcher plusieurs personnes d'utiliser la même identité [2], [16].

Mode identification : En mode identification, le système reconnaît un individu en recherchant une correspondance dans les modèles de tous les utilisateurs de la base de données. Par conséquent, le système effectue une comparaison un-à-plusieurs pour établir l'identité d'un individu (ou échoue si l’individu n'est pas inscrit dans la base de données du système) sans que l’individu ait à revendiquer une identité (par exemple, "À qui sont ces données biométriques ?"). L'identification est un élément essentiel dans les applications de reconnaissance négative, où le système établit si la personne est celle qui nie être. Le but de la reconnaissance négative est d'empêcher qu'une même personne utilise plusieurs identités.

L'identification peut également être utilisée en reconnaissance positive pour plus de commodité (l'utilisateur n'est pas tenu de revendiquer une identité) [2], [16].

I.4 Les limites des systèmes biométriques unimodaux

Concernant les systèmes unimodaux, chaque modalité en soi ne peut pas toujours être utilisée de manière fiable pour faire la reconnaissance dû aux limitations et problèmes critiques qui peuvent affecter de manière significative la fiabilité et les performances de ces systèmes. Ces limitations et problèmes peuvent être divisés en trois types principaux : précision, évolutivité, sécurité et confidentialité [2], [3], [17].

I.4.1 Précision

Idéalement, tout système biométrique devrait pouvoir offrir un haut niveau de précision dans la reconnaissance de l'identité d'un individu. Cependant, la précision des systèmes biométriques qui fonctionnent sur un seul trait est généralement affectée par un certain nombre de facteurs qui peuvent être résumés comme suit :

Bruit introduit par le capteur : il existe un certain nombre de raisons qui conduisent à l'apparition de bruit dans les données numérisées; certaines de ces raisons sont les conditions environnementales qui entourent le processus d'acquisition biométrique. Par exemple, lorsque des données vocales sont capturées dans un environnement bruyant ou que l'image du visage d'un utilisateur est capturée dans de mauvaises conditions d'éclairage, cela affecte la précision du système biométrique. De plus, les capteurs doivent être correctement entretenus. Par exemple, la qualité du trait d'empreinte digitale pourrait être faible en raison de l'accumulation de saleté sur la surface du lecteur d'empreintes digitales [18]. La mauvaise qualité des données biométriques peut entraîner une réduction significative de la précision globale du système.

Non-universalité : signifie simplement qu'un sous-ensemble de la population n'est pas en mesure de fournir correctement la modalité biométrique requise, ou ne possède pas une modalité biométrique spécifique. Par exemple, les personnes ayant de longs cils ou des anomalies oculaires peuvent ne pas être en mesure de fournir correctement leur trait d'iris à un système de reconnaissance de l'iris. Un autre exemple, où les utilisateurs ayant les doigts très secs ou gras et les travailleurs souffrant d'ecchymoses et de coupures ont des empreintes digitales qui affectent et limitent les performances d'un système de reconnaissance d'empreintes digitales [2].

Variations intra-classes : Fait référence aux variations entre les échantillons du même utilisateur acquis lors des phases d'enrôlement et de reconnaissance. Ces variations peuvent être le résultat d'un fonctionnement incorrect du capteur (par exemple, rotation et pression sur le capteur d'empreintes digitales), de changements inhérents tels que, les cicatrices et les ecchymoses dans le trait d'empreinte digitale ; d'un fraudeur imitant un trait de comportement particulier, tel que la voix , l'effet du vieillissement sur l'apparence du visage, l'effet de la maladie sur le trait de l'iris et l'utilisation de différents capteurs ou de paramètres de capteur modifiés lors des phases d'enrôlement et de reconnaissance. Dans un système biométrique idéal, les caractéristiques extraites du trait biométrique doivent être relativement invariantes à ces variations. Cependant, les caractéristiques sont très sensibles à ces variations dans la plupart des cas et peuvent nécessiter des algorithmes de classification complexes pour gérer ces variations [19].

I.4.2 Évolutivité

Ce problème concerne essentiellement l'augmentation du nombre d'utilisateurs inscrits dans la base de données et son effet sur la vitesse et les performances du système. En réalité, cet effet est différent selon la tâche fournie par le système biométrique (soit la vérification ou l'identification). Dans la tache de vérification, le système n'est pas affecté pas avec l'augmentation du nombre d'utilisateurs inscrits, car seule une comparaison un-à-un est nécessaire pour vérifier l'identité de l'utilisateur. Au contraire, l'augmentation du nombre d'utilisateurs inscrits dans un système d'identification a de sérieux effets négatifs, car le modèle de requête sera comparé à N modèles stockés dans la base de données (une comparaison un-à-plusieurs) [20], cela peut diminuera le débit du système biométrique en augmentant le temps écoulé nécessaire pour établir l'identité de l'utilisateur, et peut diminuer la précision de la reconnaissance. Pour éliminer ou réduire cet effet sur le système d'identification, une question doit être résolue : « comment pouvons-nous réduire le nombre de ces comparaisons requises ? ». En fait, certaines approches travaillent sur l'indexation ou le filtrage des utilisateurs inscrits en fonction de facteurs extrinsèques (par exemple, l'âge, le sexe, la race, etc.) ou de facteurs intrinsèques par exemple, les principales classes d'empreintes digitales. Par conséquent, seul un sous-ensemble de la base de données complète sera nécessaire à des fins de comparaison, mais ces approches ont encore certaines limites [21].

I.4.3 Sécurité et confidentialité

Ceci est généralement lié à des problèmes d'usurpation biométrique et à la possibilité que le système unimodal puisse être trompé, par exemple en utilisant une fausse empreinte digitale ou un faux modèle d'iris. Les modalités comportementales telles que la voix et la démarche sont plus vulnérables à de telles attaques que les modalités physiologiques [3]. L'un des défis les plus importants des systèmes biométriques est de savoir comment protéger tous les modèles des utilisateurs, qui sont stockés dans la base de données, contre le vol. Il existe différentes techniques pour gérer les problèmes d'usurpation biométrique, telles que les techniques de détection de la vivacité [22] pour les modalités physiques et un mécanisme de défi-réponse [3] pour les modalités physiques et comportementales.

I.5 La biométrie multimodale

Au cours des dernières années, la plupart des systèmes biométriques qui ont été largement utilisés dans des applications gouvernementales et civiles sensibles ont été des systèmes biométriques unimodaux. Bien que ceux-ci puissent fournir un haut niveau de sécurité dans la reconnaissance de l'identité d'une personne, les limitations et les problèmes mentionnés précédemment nous obligent à se déplacer vers les systèmes biométriques multimodaux.

I.5.1 Pourquoi la multimodalité?

Les systèmes multimodaux sont des systèmes qui fusionnent plusieurs sources de caractéristiques biométriques afin d'améliorer la sécurité et d'augmenter la fiabilité dans l'établissement de l'identité de la personne. En d'autres termes, les systèmes multimodaux sont principalement basés sur les données présentées et acquises à partir de multiples ressources, comme cela sera expliqué plus loin. Récemment, les systèmes multimodaux ont été adoptés pour de nombreuses applications gouvernementales et civiles en raison de leur capacité à traiter et à surmonter certains des principaux inconvénients et limitations présentés par les systèmes biométriques unimodaux. Certains des avantages des systèmes multimodaux par rapport aux systèmes biométriques unimodaux sont énumérés ci-dessous [3], [17]:

1. Augmenter l'efficacité et la fiabilité du système de reconnaissance en réduisant considérablement l'effet de bruit ou de mauvaise qualité dans les modalités biométriques établis. La disponibilité d'autres sources biométriques dans le système multimodal peut améliorer les performances du système, par exemple, si un utilisateur ne peut pas être identifié par sa voix en

raison de problèmes environnementaux, il peut toujours être identifié à l'aide d'une autre modalité, tel qu'une empreinte digitale.

2. Une couverture suffisante de la population peut être obtenue à l'aide de systèmes multimodaux, en raison de leur capacité à résoudre les problèmes liés à la phase d'enrôlement, tels que la non-universalité. Ainsi, une personne qui ne peut pas présenter un trait biométrique particulier peut toujours être enrôlée et identifiée en présentant un autre trait biométrique, par exemple, un travailleur manuel qui a une mauvaise qualité d'empreintes digitales, peut toujours être enrôlé et identifié à l'aide de traits tels que le visage, l'iris, voix, etc.

3. Un système multimodal peut réduire considérablement le chevauchement entre les espaces de caractéristiques de différentes personnes (similitudes inter-classes) en combinant des traits biométriques et en adoptant une approche de fusion. La combinaison de plusieurs éléments de preuve provenant de différentes sources peut augmenter la dimensionnalité du vecteur de caractéristiques, mais la précision globale du système biométrique sera augmentée. Par exemple, deux jumeaux qui peuvent avoir le même trait de voix et visage n'auront pas les mêmes traits d'iris et d'empreintes digitales [19].

4. Les systèmes multi-biométriques peuvent fournir une plus grande précision et une plus grande résistance à l'accès non autorisé par un imposteur que les systèmes biométriques unimodaux, en raison de la difficulté d'usurper ou de falsifier plusieurs traits biométriques pour un utilisateur légitime en même temps. De plus, un autre mécanisme peut être couplé au système multi-biométrique, par exemple en demandant aux utilisateurs de présenter leur trait biométrique de manière aléatoire au moment de l'acquisition (par exemple, empreinte digitale, suivi du visage puis de la voix) pour s'assurer que c'est un réel utilisateur en ligne qui est en interaction avec le système. C'est ce qu'on appelle la détection de l'état de présence ou le mécanisme de défi-réponse [19].

5. L'utilisation d'un système multi-biométrique peut améliorer considérablement le temps de traitement d'un système biométrique, en particulier dans la tâche d'identification où une comparaison un-à-plusieurs est nécessaire. Cette comparaison peut être effectuée en utilisant le trait biométrique le plus rapide (par exemple, le trait d'empreinte digitale) pour réduire la taille de la base de données à une taille acceptable, puis le trait biométrique le plus précis (plus lent ; par exemple, le trait d'iris) sera utilisé sur le reste de la base de données pour prendre la décision finale.

6. Enfin, une grande flexibilité sera offerte à l'utilisateur pendant le temps de reconnaissance grâce à un système multi-biométrique. Supposons qu'un système soit mis en œuvre à l'aide de trois caractéristiques biométriques (par exemple, le visage, les empreintes digitales et la voix). Par la suite, lors de la phase de reconnaissance, un utilisateur peut choisir de fournir un sous-ensemble ou l'ensemble de ses traits biométriques selon la nature de l'application mise en œuvre et la convenance de l'utilisateur.

I.5.2 Les différents systèmes multimodaux

Il existe plusieurs scénarios de fusion de traits biométriques qui dépendent essentiellement du type de sources et des caractéristiques utilisées (cf. Figure I.4) [11].

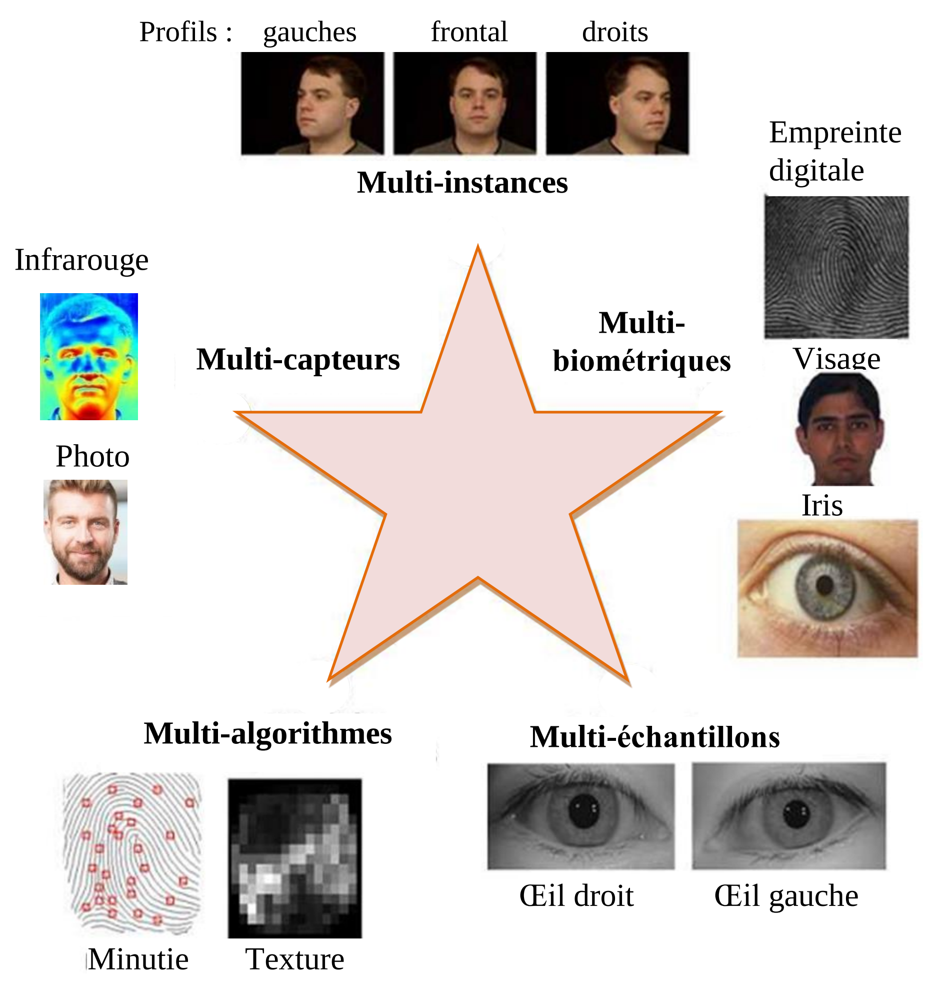


Figure I.4 Les différents types des systèmes multimodaux

1. Systèmes multi-capteurs : Un système multi-capteurs permet d'obtenir des images de la même modalité biométrique en utilisant plus d'un capteur, par exemple une caméra photographique et un capteur infrarouge de profondeur pour l'acquisition du visage en 3D. L'incorporation de plusieurs capteurs entraîne naturellement une augmentation du coût du système, mais cette approche peut aider à résoudre un biais ou une lacune particulière dans un capteur spécifique en obtenant un consensus entre les capteurs.

2. Systèmes multi-instances : lorsqu'ils associent plusieurs instances de la même biométrie, par exemple l'acquisition de plusieurs images de visage avec des changements de pose, d'expression ou d'illumination.

3. Systèmes multi-échantillons : lorsqu'ils associent plusieurs échantillons différents de la même modalité par le même capteur, dans le but de prendre les complètes variations qui peuvent se produire au sein de cette modalité, par exemple deux empreintes digitales de doigts différents ou les deux iris. Dans ce cas les données sont traitées par le même algorithme mais nécessitent des références différentes à l'enregistrement contrairement aux systèmes multi-instances qui ne nécessitent qu'une seule référence.

4. Systèmes multi-algorithmes : lorsque plusieurs algorithmes traitent la même image acquise de la même modalité, une multiplicité d'algorithmes peut intervenir dans le module d'extraction en considérant plusieurs ensembles de caractéristiques, et/ou dans le module de comparaison en utilisant plusieurs algorithmes de comparaison qui peuvent améliorer la performance du système.

5. Systèmes multi-biométriques : lorsque l'on considère plusieurs modalités biométriques différentes, par exemple le visage et l'iris. C'est le sens le plus classique du terme multimodal. Cette combinaison fournit une nette amélioration de la performance d'un système. Ces systèmes nécessitent différents capteurs ainsi que des algorithmes dédiés à chaque modalité biométrique.

6. Systèmes hybrides : Un système multimodal peut bien sûr combiner ces différents types d'associations, par exemple l'utilisation du visage et de l'iris mais en utilisant plusieurs algorithmes**.**

Tous ces types de systèmes peuvent pallier à des problèmes différents et ont chacun leurs avantages et inconvénients. Les quatre premiers systèmes combinent des informations

issues d'une seule et même modalité ce qui ne permet pas de traiter le problème de la non-universalité de certaines biométries ainsi que la résistance aux fraudes, contrairement aux systèmes "multi-biométries".

I.5.3 Modes de fonctionnement

Après avoir déterminé les sources biométriques à intégrer, l'étape suivante consiste à construire l'architecture du système. Un système multimodal peut fonctionner selon l'un des trois modes suivants : mode série, mode parallèle ou mode hiérarchique (Figure I.5) [23].

1. Mode série : dans le mode de fonctionnement en série, la sortie d'une modalité est généralement utilisée pour réduire le nombre d'identités possibles avant que la modalité suivante ne soit utilisée [23]. Par conséquent, il n'est pas nécessaire d'acquérir simultanément plusieurs sources d'information (par exemple, plusieurs modalités). En outre, une décision peut être prise avant l'acquisition de tous les traits, cela peut réduire le temps global de reconnaissance.

2. Mode parallèle : dans le mode de fonctionnement parallèle, les informations provenant de plusieurs modalités sont utilisées simultanément et indépendamment. Ensuite, les résultats sont combinés pour prendre la décision finale de classification. Par exemple, un système d'authentification basé sur la reconnaissance des visages et d'iris.

3. Mode hiérarchique : dans ce mode, les classificateurs individuels sont combinés dans une structure arborescente, ce mode est pertinent lorsque le nombre de classificateurs est important.

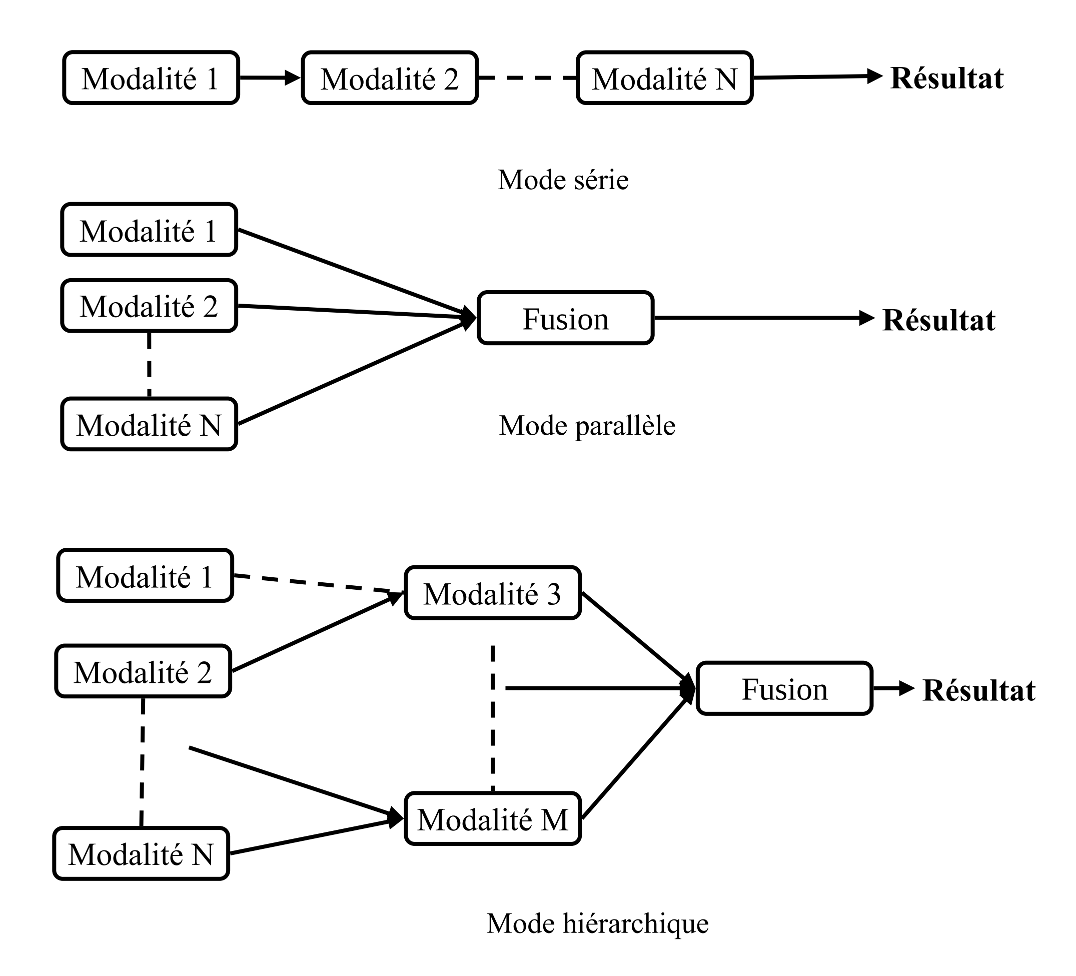


Figure I.5 Les trois modes de fonctionnement d'un system multimodal

I.5.4 Les différents niveaux de fusion

Dans les systèmes multimodaux, le terme fusion est souvent utilisé pour décrire le processus de combinaison d'informations provenant de plusieurs sources dans le processus de reconnaissance. La section précédente a décrit les différents types des systèmes multimodaux qui peuvent être utilisés pour améliorer les performances; la fusion est utilisée pour combiner les résultats de la redondance afin qu'une seule sortie puisse être produite à partir de toutes les modalités utilisées. Il existe quatre niveaux auxquels la fusion peut se produire dans un système multimodal (Figure I.6).

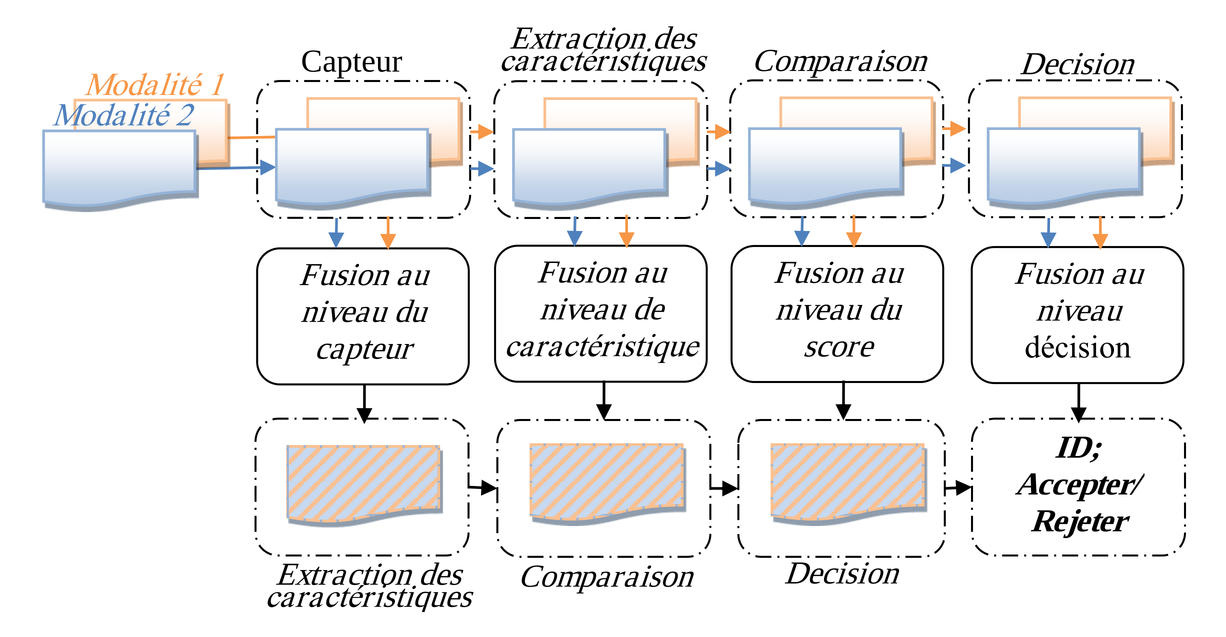


Figure I.6 Les différents niveaux de fusion

1. Fusion au niveau du capteur : À ce niveau de fusion, les sources d'informations utilisent différents capteurs compatibles pour capturer la même modalité biométrique ou capturent plusieurs instances de la même modalité à l'aide d'un seul capteur [24]. Par exemple, le visage peut être obtenu en utilisant deux capteurs différents, une caméra photographique pour les couleurs et les détails, et une caméra infrarouge pour la profondeur, cette fusion produit un visage en 3D (exemple de capteur, Microsoft Kinect) [19] .

2. Fusion au niveau caractéristique : La forme la plus courante de la fusion au niveau des caractéristiques dans la littérature est la concaténation de vecteurs de caractéristiques. Dans cette approche, des vecteurs de caractéristiques sont extraits pour chaque modalité, et sont ensuite concaténés pour former un vecteur de caractéristiques commun. Souvent, ce processus est suivi d'un processus de sélection et/ou de normalisation dans lequel l'espace des caractéristiques est réduit à un ensemble uniforme et plus compact [3].

3. Fusion au niveau score: À l'aide de la fusion au niveau des scores, les scores de similarité ou de distance renvoyés par chaque comparateur sont combinés pour produire un score de sortie final. Étant donné que les sorties de divers algorithmes de comparaison peut se situer dans différents domaines, une étape de normalisation des scores est souvent nécessaire, dans laquelle les distributions des scores de différents comparateurs sont modifiées dans la même échelle pour être compatibles avec les étapes de fusion ultérieures [25].

4. Fusion au niveau décision : La fusion au niveau de la décision applique un comparateur à chaque échantillon biométrique (ou le même comparateur à plusieurs échantillons) pour obtenir une réponse booléenne indiquant si chaque comparaison est une correspondance ou non. Les sorties sont ensuite fusionnées à l'aide d'opérateurs booléens, d'un schéma de vote ou d'une méthode de probabilité [26].

I.6 L'apprentissage approfondi (Deep Learning)

L'apprentissage approfondi, qui est une branche de l'intelligence artificielle, aujourd'hui il considéré comme une technologie de base de la quatrième révolution industrielle [27]. En raison de ses capacités d'apprentissage, il devenue un sujet important dans le domaine de la technologie, il est largement utilisé dans diverses applications tels que les soins de santé, la reconnaissance visuelle, l'analyse de texte, la cyber sécurité et bien d'autres. Dans ce parti, nous allons découvert les algorithmes d'apprentissage approfondi les plus utilisés.

I.6.1 Perceptron multicouche (MLP)

Il est également connu comme l'architecture de base de l'apprentissage approfondi. Un MLP standard est un réseau entièrement connecté qui se compose d'une couche d'entrée qui reçoit les données d'entrée, d'une couche de sortie qui prend une décision ou fait une prédiction sur le signal d'entrée, et d'une ou plusieurs couches cachées entre ces deux couches [28], [29]; Figure I.7 montre un exemple d'un MLP. La sortie d'un réseau MLP est déterminée à l'aide d'une variété de fonctions d'activation, telles que ReLU (Rectifed Linear Unit), Tanh, Sigmoïde et Softmax [30]. Pour entraîner le MLP, on utilise l'algorithme de Rétropropagation [28], qui est également connue comme étant l'élément de base d'un réseau de neurones. Le MLP nécessite d'avance le paramétrage du nombre de couches cachées, de neurones et d'itérations, ce qui peut facilite la résolution d'un modèle compliqué. Cependant, grâce à l'ajustement partiel, le MLP offre l'avantage d'apprendre des modèles non linéaires en temps réel ou en ligne [30].

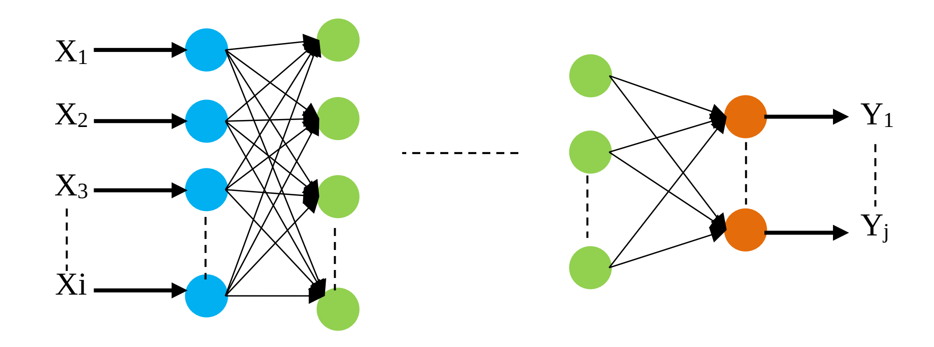


Figure I.7 Le schéma générale d'un Perceptron multicouche

I.6.2 Réseau de neurones convolutifs (CNN)

Le réseau de neurones à convolution (CNN ou ConvNet) [31] est une architecture d'apprentissage approfondi discriminatif populaire, qui apprend directement à partir de l'entrée sans nécessiter d'extraction de caractéristiques. La Figure I.8 montre un exemple de CNN comprenant plusieurs couches. Par conséquent, le CNN améliore la conception des réseaux de neurones traditionnels tels que les réseaux MLP régularisés. Chaque couche du CNN prend en compte les paramètres optimaux pour un résultat significatif et réduit la complexité du modèle. Le CNN utilise également un "dropout" [32] qui peut traiter le problème de sur-ajustement, qui peut se produire dans un réseau traditionnel.

Les CNN sont spécifiquement conçus pour traiter une variété de formes 2D, donc ils sont largement utilisés dans la reconnaissance visuelle, l'analyse d'images médicales, la segmentation d'images, le traitement du langage naturel, et bien d'autres encore [27]. La capacité à découvrir automatiquement les caractéristiques essentielles des données d'entrée sans intervention humaine les rend plus puissants qu'un réseau traditionnel. Plusieurs variantes de CNN existent dans le domaine, notamment le groupe de géométrie visuelle (VGG) [33], AlexNet [34], ResNet [35], etc. qui peuvent être utilisées dans divers domaines d'application en fonction de leurs capacités d'apprentissage.

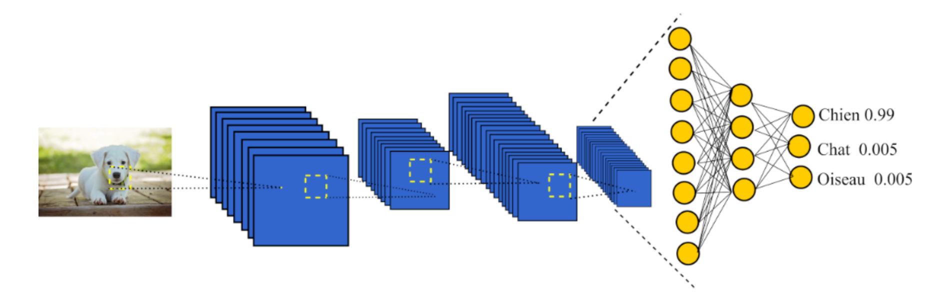


Figure I.8 Le schéma générale d'un CNN

I.6.3 Machine de Boltzmann restreinte (RBM)

Une machine de Boltzmann restreinte (RBM) [36] est également un réseau neuronal stochastique génératif capable d'apprendre une distribution de probabilité sur ses entrées. Les machines de Boltzmann se composent généralement de nœuds visibles et cachés et chaque nœud est connecté à tous les autres, ce qui nous aide à comprendre les irrégularités en apprenant comment le système fonctionne dans des circonstances normales [27]. Les RBM sont un sous-ensemble des machines de Boltzmann qui limitent le nombre de connexions entre les couches visibles et cachées. Cette restriction permet aux algorithmes d'apprentissage tels que l'algorithme de divergence contrastive basé sur le gradient d'être plus efficaces que ceux des machines de Boltzmann en général [37].

Les RBM ont trouvé des applications dans la réduction de la dimensionnalité, la classification, la régression, le filtrage collaboratif, l'apprentissage de caractéristiques, et bien d'autres. Dans le domaine de la modélisation de l'apprentissage approfondi, ils peuvent être entraînés de manière supervisée ou non supervisée, en fonction de la tâche.

I.6.4 Encodeur automatique (Auto-Encoder: AE)

Un Autoencodeur (AE) [38] est une technique populaire d'apprentissage dans laquelle les réseaux neuronaux sont utilisés pour apprendre des représentations. Généralement, les AE sont utilisés pour travailler avec des données à haute dimension, et la réduction de la dimensionnalité explique comment un ensemble de données est représenté. Le codeur, le code et le décodeur sont les trois parties d'un AE. L'encodeur compresse l'entrée et génère le code, que le décodeur utilise ensuite pour reconstruire l'entrée [27]. L'AE est largement utilisé dans de nombreuses tâches d'apprentissage, par exemple la réduction de la dimensionnalité, l'extraction de caractéristiques, le codage efficace, la modélisation générative, l'effacement de bruit, etc. [39]. Les Spare Autoencodeur (SAE) [40], comme nous le verrons plus loin en détail dans le chapitre II.4.7 peuvent avoir plus d'unités cachées que d'entrées, mais seul un petit nombre d'unités cachées est autorisé à être actif en même temps, ce qui donne un modèle spare. Ce modèle est donc obligé de répondre aux caractéristiques statistiques uniques des données d'apprentissage en suivant ses contraintes.

I.7 Conclusion

Dans ce chapitre, nous avons définit la biométrie d'une manière générale, et décrit les différents modules qui constituent un système pour la reconnaissance biométrique. Nous avons présenté également les limites des systèmes biométriques unimodaux, et comment la multimodalité biométrique fonctionne pour résoudre ces limitations.

Chapitre II

Reconnaissance biométrique par le visage

II.1 Introduction

La reconnaissance faciale a suscité beaucoup d'intérêt au cours des dernières décennies, notamment en raison de ses énormes domaines d'application, telle que la vidéosurveillance, la sécurité publique, et l'interaction homme-machine [22]–[24]. Le rôle principal de la reconnaissance faciale est d'identifier une image de visage à partir d'un ensemble de plusieurs images des visages stockées dans une base de données [44]. Les principales méthodes utilisées pour la reconnaissance faciale sont appelées méthodes holistiques, y compris : l'analyse en composantes principales (PCA) [45], kernel PCA (KPCA) [46], l'analyse discriminante linéaire (LDA) [47], et kernel LDA [48]. Sur la base de PCA et LDA, Mandal (2009) a introduit l'application de la transformation curvelet en conjonction avec PCA-LDA [49], Huang (2010) a fusionné à la fois les informations entre les lignes et les colonnes à l'aide de 2DPCA et 2DLDA sur les images de visage [50] ; Wen (2012) a proposé une approche pour la reconnaissance de visage basée sur les vecteurs de différence et le KPCA (DV-KPCA) [51] ; Huang(2014) a proposé une structure locale préservant l'analyse discriminante (LSPDA) [52].

Cependant, les méthodes globales basées sur PCA et LDA, également connues sous le nom de eigenfaces et fisherfaces, sont statistiquement linéaires et ne sont pas efficaces dans des environnements non contrôlés. En effet, les images de visage peuvent être interprétées comme des objets non linéaires [53], et les performances de ces méthodes se dégradent significativement, notamment pour les images de visage multi-vues [54], ou sous des changements d'éclairage, d’expressions, de l'âge, des occlusions et les différentes poses [55]. En conséquence, de nombreux chercheurs ont proposé d'autres techniques locales d'extraction de caractéristiques pour atténuer les limitations ci-dessus. Parmi ces techniques : local binary pattern (LBP) [56], scale-invariant feature transform (SIFT) [57], filtres de Gabor [54], [58], local phase quantisation (LPQ) [59], binarised statistical image features (BSIF) [60], et le histograms of oriented gradients (HOG) [61], ont été proposées. Malgré que ces techniques ont prouvé leur efficacité en améliorant la précision de certains problèmes de reconnaissance des visages, mais elles ont également souffert dans le cas d'une mauvaise représentation des non-linéarités de nombreuses images faciales [62].

Récemment, de nombreux succès ont été obtenus en combinant judicieusement les méthodes locales et globales, afin de fournir des informations complémentaires pour une extraction de caractéristiques plus efficace. Par exemple, Zhang (2005) a proposé la méthode

local Gabor binary pattern histogram sequence (LGBPHS) [63], qui combine Gabor et les histogrammes de LBPs pour capturer les variations d'apparence liées à l'éclairage, à l'expression et au changement de l'âge. Yu et al (2010) a proposé une représentation basée sur le module et la phase de Gabor (GMPTR), et sur la méthode null space linear discriminant analysis (NLDA) [64]. Zhou et al (2013) a décrit une approche (GLL) pour les images faciales floues et à faible résolution qui combine Gabor, LBPs et LPQs pour assurer que le flou est invariant et capable de capturer l'information de texture [65]. Yu et al (2014) a proposé d'intégrer la moyenne et l'écart-type de la différence absolue locale à partir de LBP dans l'étape de l'extraction de caractéristiques, afin d'améliorer la capacité de classification des caractéristiques [66]. Sur la base de la méthode 2D-DWT, Huang et al (2015) a proposée l'approche TWSBF pour combiner les caractéristiques et les pixels par l'emploi de la décomposition en sous-bandes d'ondelettes, et d'utilisé PCA et LDA pour réduire la dimension des caractéristiques [67]. Guermoui et Mekhalfi (2016) a proposé une représentation sparse de la méthode complete local binary pattern histogram (CLBP_H) utilisé une classification basée sur la représentation sparse et la concaténation de complete LBP sing histogram (CLBP SH) et complete LBP magnitude histogram (CLBP MH) après la représentation pyramidale pour créer des caractéristiques plus globales [68]. Fathi (2016) combine les caractéristiques global-Gabor-Zernike (GGZ) et l'histogram of oriented gradient (HOG) [69]. Cependant, ces méthodes ne sont pas supervisées ou basées sur des filtres prédéfinis, les réseaux de neurones profonds comme les réseaux de neurones convolutifs (CNN) et les modèles d'Autoencodeurs (AE) [38] ont récemment été fournis pour l'apprentissage direct des filtres liés au problème étudié [38], [70]–[72]. Peng et al(2015) a proposé la méthode discriminative graph regularised extreme learning machine (GELM) pour améliorer encore les performances de classification du modèle de réseaux neuronaux «extreme learning machine (ELM) »[73]. Liu et al (2018) proposé un modèle d'apprentissage approfondi appelé enhanced PCA network (EPCANet) sur la base de CNN pour la classification de visage a utilisé deux couches de convolution pour l'apprentissage des filtres PCA [74].

L'apprentissage en profondeur nécessite généralement un traitement intensif, et de grandes ressources pour l'apprentissage des modèles, afin de réduire les distorsions, en particulier dans des environnements non contrôlés [75]. Ces conditions peuvent ne pas être réalisables dans les systèmes réels de reconnaissance des visages, dans ce cas, ce chapitre propose une nouvelle méthode de reconnaissance des visages, en utilisant une combinaison de

caractéristiques et d'apprentissage approfondi par l'Autoencodeurs (AE). Au lieu d'utiliser les images brutes du visage dans leur espace d'origine (images 2D) comme décrit dans la littérature, nous proposons une combinaison de plusieurs caractéristiques après la réduction comme une entrée de l'Autoencodeur. Nous proposons une combinaison de plusieurs caractéristiques après une étape de prétraitement utilisant la méthode adaptative de l'égalisation d'histogramme limitée par le contraste (CLAHE) [76], pour créer un nouvel espace de représentation pour l'entraînement de l'AE. Les caractéristiques utilisées dans ce travail sont une combinaison des caractéristiques locales et globales en utilisant Gabor, LBP, et LPQ pour assurer une représentation plus discriminante, puis que le système est invariant dans des environnements non contrôlés. Après une réduction de chaque vecteur caractéristique indépendamment en utilisant le LDA, le AE est entraîné par les trois caractéristiques fusionnées et réduites après un processus de normalisation.

Ce chapitre présente notre système proposé de la reconnaissance biométrique par le visage. L'objectif de ce chapitre est d'explorer les développements de notre système de reconnaissance biométrique sur le visage et les algorithmes qui les sous-tendent.

II.2 Applications

La reconnaissance des visages est utilisée pour deux tâches principales :

1. **Vérification** (comparaison un-à-un) : Lorsque l'on présente l'image du visage d'un individu inconnu accompagnée d'une déclaration d'identité, il s'agit de vérifier si l'individu est bien celui qu'il prétend être.
2. **Identification (**comparaison un-à-plusieurs) : Étant donné l'image d'un individu inconnu, déterminer l'identité de cette personne en comparant cette image (caractéristiques) avec une base de données d'images d'individus connus.

Il existe de nombreux domaines d'application dans lesquels la reconnaissance des visages peut être exploitée à ces deux tâches, dont quelques-uns sont présentés ci-dessous.

*** Sécurité** : contrôle d'accès aux bâtiments, aux aéroports/ports, aux distributeurs automatiques de billets (ATM), et aux postes de contrôle frontaliers; sécurité des ordinateurs/réseaux.

*** Surveillance :** un grand nombre de caméras de vidéo-surveillance peuvent être contrôlées pour rechercher des criminels connus, des trafiquants de drogue, …etc. et les autorités peuvent être informées lorsque l'un d'entre eux est localisé.

*** Vérification générale de l'identité :** enregistrement électoral, opérations bancaires, commerce électronique, cartes d'identité nationales, passeports, permis de conduire, cartes d'identité des employés.

*** Systèmes de justice criminelle :** (systèmes de photos d'identité judiciaire, analyse post-événement, médecine légale).

*** Les applications de type "carte à puce" :** au lieu de maintenir une base de données d'images faciales, l'empreinte du visage peut être stockée dans une carte à puce, un code à barres ou une bande magnétique, dont l'authentification est effectuée en faisant correspondre l'image réelle et le modèle stocké [77].

*** Enquêtes sur les bases de données d'images :** recherche dans les bases de données d'images de conducteurs titulaires d'un permis de conduire, de bénéficiaires de prestations, d'enfants disparus, d'immigrants et d'enregistrements de la police).

*** Environnements multimédias avec interfaces homme-machine adaptatives :** (partie de systèmes ubiquitaires ou contextuels, surveillance du comportement dans les crèches ou les centres pour personnes âgées, reconnaissance d'un client et évaluation de ses besoins) [78].

*** Indexation vidéo :** étiquetage des visages dans la vidéo [79], [68].

En plus de ces applications, les techniques utilisées dans la reconnaissance des visages ont également été modifiées et utilisées pour des applications connexes telles que la classification des genres [80], [81], la reconnaissance des expressions [82], [83] et la reconnaissance et le suivi des caractéristiques faciales [84] ; chacune de ces applications a son utilité dans divers domaines : par exemple, la reconnaissance des expressions peut être utilisée dans le domaine de la médecine pour le suivi des soins intensifs, tandis que la reconnaissance et la détection des caractéristiques faciales peuvent être exploitées pour suivre les yeux d'un conducteur de véhicule et ainsi surveiller sa fatigue [85], ainsi que pour la détection du stress [86]. La reconnaissance faciale est également utilisée en conjonction avec d'autres méthodes biométriques telles que la reconnaissance de la parole, de l'iris, des empreintes digitales, de

l'oreille et de la démarche, afin d'améliorer les performances de reconnaissance de ces méthodes [6], [7], [13], [15], [17], [20], [21].

II.3 Les Difficultés dans la reconnaissance du visage

La reconnaissance des visages est un cas spécifique et difficile de la reconnaissance des formes dans le cas général, et en biométrie spécifiquement. La difficulté de ce problème provient du fait que les visages semblent être à peu près semblables et les différences entre eux sont assez subtiles. Par conséquent, les images de visages frontaux forment un groupe très dense dans l'espace image, ce qui rend pratiquement impossible pour les techniques traditionnelles de reconnaissance des formes de les distinguer avec un haut degré de réussite [89]. Ainsi que les difficultés associées à l'environnement, qui pourrait devenir non contrôlé dans le cas réel, ces difficultés peuvent être attribuées aux facteurs suivants [90]:

1. **Intensité** : Il existe trois types d'intensité : couleur, gris et binaire.
2. **Pose** : Les images de visages varient en fonction de la pose relative de la tête (frontale, 45°, profil), et certains traits du visage, comme un œil peuvent être partiellement ou totalement occultés.
3. **Composants structurels :** Les traits du visage tels que la barbe et la moustache peuvent être présents ou non.
4. **Rotation de l'image :** Les images de visages varient directement pour différentes rotations.
5. **Mauvaise qualité :** En cas d'images de mauvaise qualité, par exemple des images floues, des images déformées et des images avec du bruit, l'intensité de l'image devient inhabituelle.
6. **L'expression faciale :** L'apparence des visages dépend d'une expression faciale personnelle.
7. **Changement de l'âge :** Les visages humains toujours en changement, grâce à l'âge et la nature.
8. **Occlusion :** Les visages peuvent être partiellement occultés par d'autres objets tels que la main, le foulard, les lunettes, etc.
9. **Illumination :** Les images des visages varient en fonction de la luminosité.

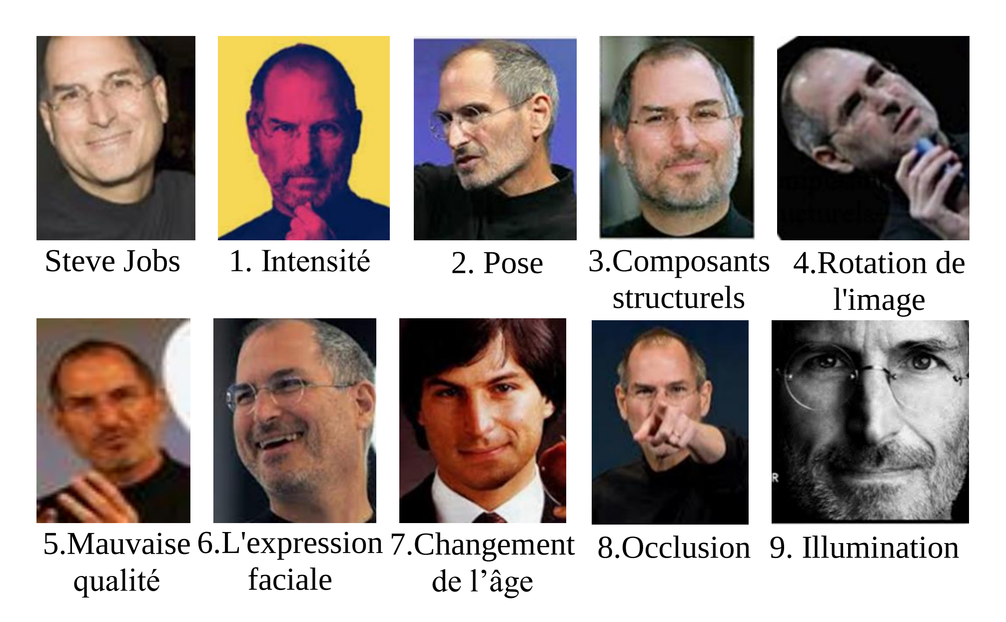


Figure II.1 Exemple des difficultés dans la reconnaissance du visage

II.4 Le système proposé

La méthode Feature Autoencodeurs (FAE) proposé, inclut les étapes de prétraitement, d'extraction de caractéristiques, de réduction, de normalisation et de modèle d'apprentissage approfondi (AE), comme le montre la Figure II.2.

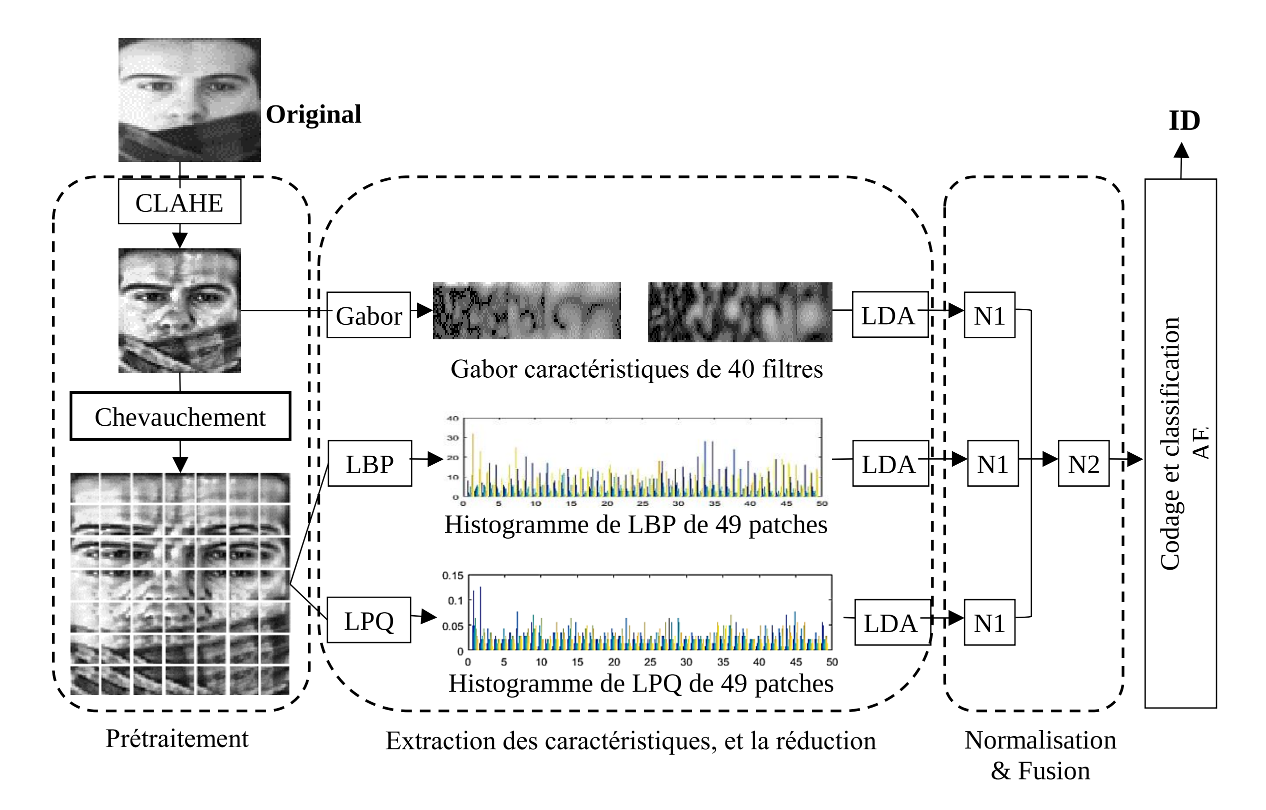


Figure II.2 L'approche FAE proposée.

II.4.1 Prétraitement

L'algorithme CLAHE [76] est utilisé pour améliorer le contraste localement sur de petites régions des images, cette amélioration est limitée par un niveau de coupure prédéfini (CL) pour réduire les problèmes d'augmentation du bruit produit par l'amélioration de contraste dans certaines conditions. Dans ce travaille, on prend CL=0.01 comme décrit par Sharma et al. [91], qui fournit la meilleure performance après de nombreuses expériences, et le meilleur contraste dans toutes les bases de données, comme le montre la Figure II.3.



Figure II.3 Résultats de l'amélioration par CLAHE sur les bases des données suivantes : (a) CMU PIE, (b) Yale, (c) AR.

Après l'étape de prétraitement, les images obtenues partitionnées en P patchs à l'aide d'une fenêtre glissante de (m × n) pixels avec un chevauchement de 50%, comme le montre la Figure II.4.

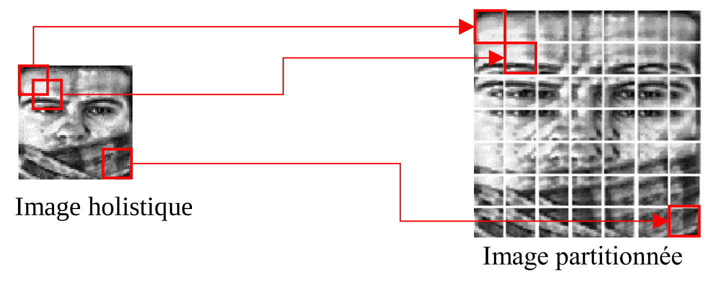


Figure II.4 Exemple de partitionnement d'une image de taille 64 × 64 divisée en 49 patchs avec un chevauchement de 50 %, où m × n = 16 × 16

Dans ce travail, nous avons choisi de combiner Gabor, LBP et LPQ comme vecteurs de caractéristiques d'entrée pour l'AE.

II.4.2 Le filtre de Gabor

Cette section présente brièvement l'utilisation des filtres de Gabor dans notre travail pour la reconnaissance des visages. Elle commence par l'introduction des filtres de Gabor et les concepts de base de l'extraction de caractéristiques à l'aide du banc de filtres de Gabor (module), qui constitue la base de nombreuses techniques populaires de reconnaissance des visages.

1. Construction d'un filtre de Gabor : Les filtres de Gabor (également appelés ondelettes ou noyaux de Gabor), ont prouvé qu'ils constituaient un outil puissant pour l'extraction des caractéristiques et la reconnaissance faciale, ils représentent des filtres complexes limités avec une localisation optimale dans le domaine spatial et fréquentiel. Ainsi, lorsqu'ils sont utilisés pour l'extraction de caractéristiques faciales, ils permettent d'extraire des caractéristiques multi-résolutionnelles et spatialement locales d'une bande de fréquence limitée [92]. En général, la famille des filtres de Gabor 2D peut être définie dans le domaine spatial en utilisant l'équation (1) [58].

$$g_{u,\theta}(x, y) = \left(\frac{Fu^2}{2\pi}\right) \times \exp\left(\frac{-Fu^2}{2\pi}(xc^2 + yc^2)\right) \times \exp\big((2\pi \times Fu \times xc)i\big) \quad (1)$$

Où pour chaque orientation θ et pour chaque échelle u,

$$xc = x\cos(\theta) + y\sin(\theta) \quad (2)$$

$$yc = -x\sin(\theta) + y\cos(\theta) \quad (3)$$

$$\text{And} \quad Fu = 0.25/\sqrt{2}^{u} \quad (4)$$

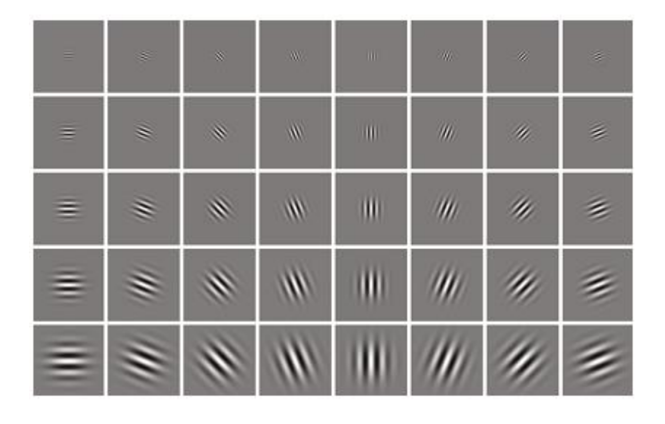


Figure II.5 Les parties réelles du banc de filtres de Gabor crée avec 5 échelles et de 8 orientations.

2. Extraction de caractéristiques à base de filtres de Gabor

Chaque image $i(x,y)$ convoluée avec 40 filtres de Gabor $g_{u,\theta}(x,y)$ crée avec $u = 5$ échelles et $\theta = 8$ orientations en utilisant l'équation (5)

$$G_{u,\theta}(x,y) = i(x,y) * g_{u,\theta}(x,y) \quad (5)$$

Où $G_{u,\theta}(x,y)$ désigne la sortie complexe du filtrage qui peut être décomposée en ses parties réelle $Re_{u,\theta}(x,y)$ et imaginaire $Im_{u,\theta}(x,y)$. Comme la majorité des techniques de reconnaissance des visages basées sur Gabor trouvées dans la littérature, qui se basent sur l'information de module lors de la construction de la représentation des visages de Gabor, dans ce travail, on prend le module"$r_{u,\theta}(x,y)$"comme une caractéristique Figure II.6.

$$r_{u,\theta}(x,y) = \sqrt{Re_{u,\theta}(x,y)^2 + Im_{u,\theta}(x,y)^2} \quad (6)$$

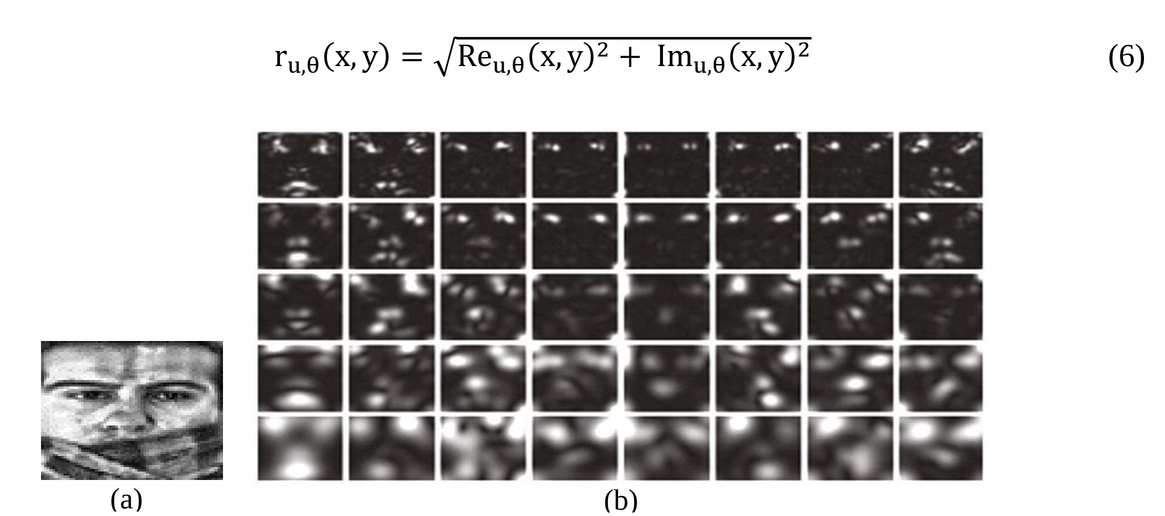


Figure II.6 Exemple de réponse de Gabor : (a) une image d'entrée, et (b) la sortie module de l'opération de filtrage avec l'ensemble du banc de 40 filtres de Gabor.

II.4.3 Motif binaire local (LBP)

L'opérateur LBP a été introduit par T.Ojala [93] pour la classification des textures. Pour un pixel central de l'image, la valeur LBP est calculée en comparant sa valeur du niveau de gris avec celle de ses (3×3) voisins (Figure II.7).

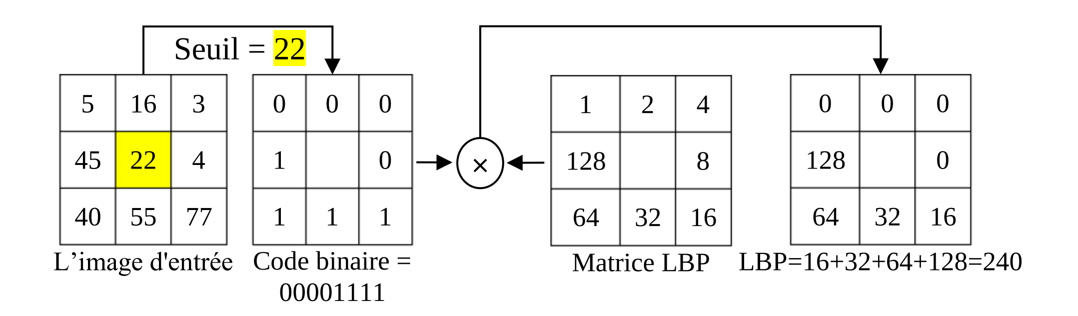


Figure II.7 L'opérateur LBP de base

Dernièrement, afin de s'adapter à la caractéristique à différentes échelles, T.Ojala [56] a étendu le voisinage 3×3 de l'opérateur LBP de base, à n'importe quelle taille, et a utilisé un voisinage circulaire au lieu d'un carré. Tous les pixels dans le voisinage circulaire de rayon R sont possibles dans l'opérateur LBP amélioré. En supposant que le pixel central est g_c , le modèle de texture LBP avec un rayon R et un nombre de pixels d'échantillonnage P construit comme montré dans la Figure II.8.

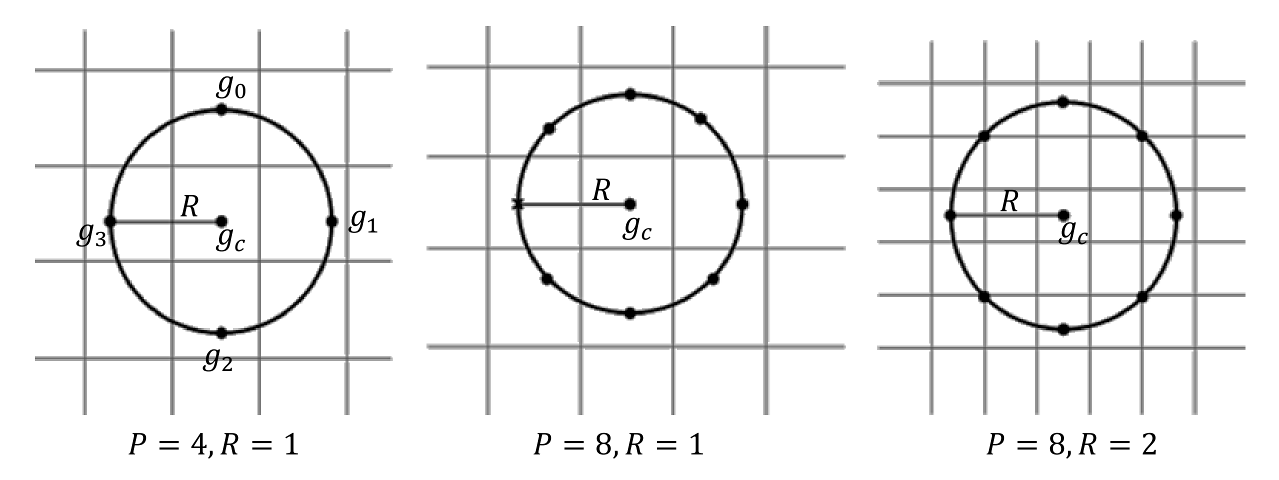


Figure II.8 Différents opérateurs LBP, pour Différents rayon ***R*** et échantillon ***P***.

La valeur d'un pixel $g_c(X_c, Y_c)$ de l'opérateur LBP circulaire est la suivante :

$$LBP_{P,R}(X_c, Y_c) = \sum_{p=0}^{P-1} S(g_p - g_c) \times 2^p \tag{7}$$

$$S(x) = \begin{cases} 1 \text{ si } x \geq 0 \\ 0 \text{ si } x < 0 \end{cases} \tag{8}$$

LBP uniforme : L'opérateur LBP produit 2^P valeurs de sortie différentes qui peuvent être formés par les P pixels de l'ensemble voisin, si tous les modèles 2^P sont adoptés, le calcul

sera très complexe. Les études montrent que certains modèles apparaissent à une faible fréquence, et certains modèles contiennent plus d'informations que d'autres. Par conséquent, il est possible d'utiliser seulement un sous-ensemble de l'ensemble 2^P de ces modèles pour décrire la texture des images. Ce type de modèles est appelé modèle uniforme [56], de la formule suivante :

$$U(LBP_{P,R}) = \left|S(g_{p-1} - g_c) - S(g_0 - g_c)\right| + \sum_{p=0}^{P-1} S(g_p - g_c) - S(g_{p-1} - g_c) \leq 2 \qquad (9)$$

Le ULBP a un point commun, où il y a deux changements de 0 à 1 au maximum dans le code binaire circulaire, par exemple, 11111111 n'a aucun changement de code, et 00111100 a deux changements de code. Le LBP avec un rayon R=1 et un point d'échantillonnage P=8 peut être exprimé comme $LBP_{8,1}$ a 256 modèles possibles, cependant, le LBP uniforme $ULBP_{8,1}$ n'a que 59 modèles possibles, ce qui réduit considérablement le calcul.

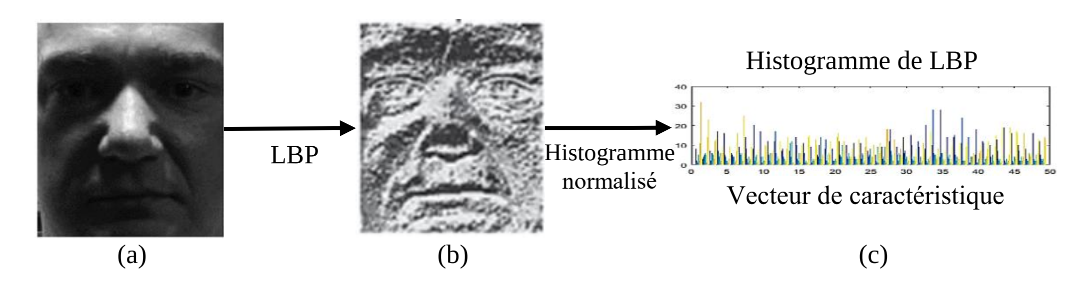


Figure II.9 Description du visage par LBP : (a) une image d'entrée, et (b) La représentation LBP, (c) vecteur de caractéristique LBP

Dans ce travail, et après plusieurs expériences, nous avons choisi l'approche de texture ULBP avec $R = 2$ et $P = 16$ pour l'extraction de l'histogramme normalisé et l'utiliser comme une vecteur de caractéristiques.

II.4.4 Quantification par phase locale (LPQ)

L'opérateur LPQ originalement créé par Ojansivu et Heikkila [59], similaire à la méthodologie LBP, LPQ est un descripteur de texture d'une image en niveaux de gris, qui est insensible au flou et invariant en rotation. L'information de phase est extraite localement sur un voisinage carré N_z de taille (M × M), en utilisant la transformée de Fourier à court terme (STFT) à chaque position de pixel z de l'image f(z) [94], [95] définie par :

$$F(\xi, z) = \sum f(z - y).\, e^{-2\pi\xi^T y} = w_\xi^T f_z \qquad (10)$$

Avec f_z c'est un vecteur contient tous les valeurs de gris M^2 de N_z, et w_ξ^T est le vecteur de base de la STFT à la fréquence ξ. Seuls quatre coefficients complexes sont considérés sur l'opérateur LPQ qui correspondent aux fréquences 2D : $\xi_1 = [a, 0]^T$, $\xi_2 = [0, a]^T$, $\xi_3 = [a, a]^T$, et $\xi_1 = [a, -a]^T$, où a est un scalaire suffisamment petit pour satisfaire $H(\xi) > 0$. L'information de phase est calculée par les points imaginaires et réels du vecteur de caractéristiques F_z. Pour chaque position de pixel z, le vecteur de caractéristiques F_z est donné par l'équation (12):

$$F_z^v = [F(\xi_1, z), F(\xi_2, z), F(\xi_3, z), F(\xi_4, z)] \quad (11)$$

$$F_z = [\mathrm{Re}\{F_z^v\}, \mathrm{Im}\{F_z^v\}] \quad (12)$$

Avec $\mathrm{Re}\{x\}, \mathrm{Im}\{x\}$ sont les parties réels et imaginaires d'un nombre complexe respectivement. La matrice de transformation $8 \times M^2$ correspondante est donnée par

$$W = \left[\mathrm{Re}\{w_{\xi_1}, w_{\xi_2}, w_{\xi_3}, w_{\xi_4}\}, \mathrm{Im}\{w_{\xi_1}, w_{\xi_2}, w_{\xi_3}, w_{\xi_4}\}\right] \quad (13)$$

$$F_z = W f_z \quad (14)$$

Avec, F_z est calculée pour la position entière de l'image. Donc, une quantification des vecteurs obtenus est effectuée avec le quantificateur scalaire :

$$q_i = \begin{cases} 1, \mathrm{si}\ f_i \geq 0 \\ 0, \mathrm{si}\ f_i < 0 \end{cases} \quad (15)$$

Où f_i est la i[ième] composante de F_z. Le codage binaire suivant est utilisé pour transformer les coefficients quantifiés en entiers (niveaux de gris) allant de 0 à 255. Ensuite, le vecteur de caractéristiques est construit avec l'histogramme normalisé de ces valeurs entières afin d'être utilisé dans les tâches de classification :

$$LPQ = \sum_{i=1}^{8} q_i \times 2^{i-1} \quad (16)$$

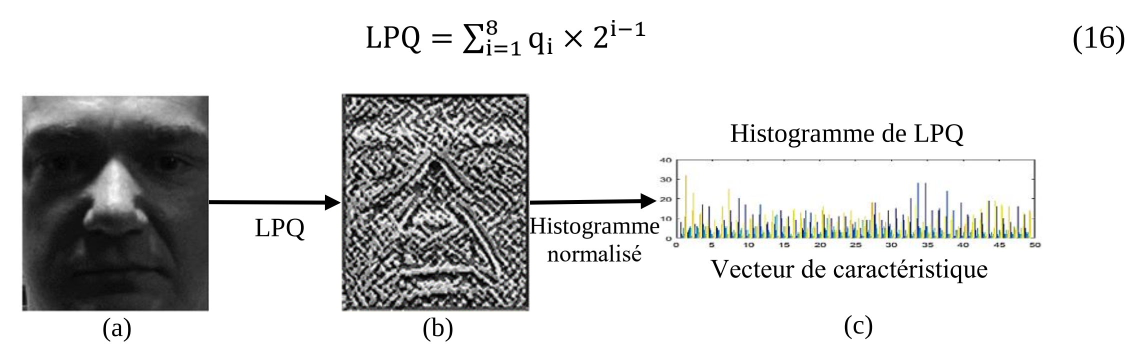


Figure II.10 Description du visage par LPQ : (a) une image d'entrée, et (b) La représentation LPQ, (c) vecteur de caractéristique LPQ

II.4.5 L'analyse discriminante linéaire (LDA)

L'analyse en discriminante linéaire (LDA) est utilisée pour réduire la dimension de l'espace initial tout en ne conservant que les propriétés les plus discriminantes des caractéristiques extraites. Plusieurs approches ont été utilisées dans la littérature pour la reconnaissance des visages afin de réduire la dimensionnalité de l'extraction des caractéristiques. Nous proposons les étapes suivantes pour discriminer les images d'entrée [96]–[99]:

Etape 1 : Nous avons besoin d'un ensemble d'entraînement composé d'un groupe important de sujets présentant diverses caractéristiques faciales. La base de données doit contenir plusieurs exemples d'images de visage pour chaque sujet dans l'ensemble d'entraînement et au moins un exemple dans l'ensemble de test. On suppose que toutes les images sont déjà normalisées en matrice $m \times n$ et qu'elles contiennent uniquement les régions du visage.

Etape 2 : Considérons un ensemble de données d'entraînement contenant N exemples$\{X_1, X_2 \ldots X_N\}$, où chaque exemple X_i est un vecteur colonne de longueur d, et appartient à l'une des K classes. Soit C_k l'ensemble de tous les exemples de la classe k, et soit $N_k = |C_k|$ le nombre d'exemples de la classe $k = 1 \ldots K$. Dans LDA, les matrices de dispersion inter-classes S_w et intra-classes S_b sont calculées par :

$$S_w = \sum_{k=1}^{K} \sum_{i=1}^{N} (X_i - m_k)(X_i - m_k)^T \tag{17}$$

$$S_b = \sum_{=1}^{K} n_k (m_k - m)(m_k - m)^T \tag{18}$$

Où, $m_k = \frac{1}{n_k} \sum_{i \in C_k} X_i$ est la moyenne de la $k^{ième}$ classe, et $m = \frac{1}{N} \sum_{i \in N} X_i$ est la moyenne de tout la base de données d'entraînement. Nous cherchons la transformation linéaire $X \longrightarrow W^T X$ qui maximise la variance inter-classes par rapport à la variance intra-classes, où W est une matrice de dimension de $d \times d$. On peut montrer que les colonnes de l'optimum W sont les vecteurs propres généralisés tels que $S_b w = \lambda S_w w$.

Dans ce travail, LDA est appliqué pour réduire les vecteurs de caractéristiques de haute dimensionnalité obtenus par les méthodes de Gabor, LBP et LPQ indépendamment.

II.4.6 La normalisation

Les trois vecteurs de caractéristiques (Gabor-LDA, ULBP-LDA, et LPQ-LDA) sont ensuite normalisés en appliquant la normalisation de la variance unitaire moyenne sur chaque vecteur de caractéristiques.

$$y = \frac{x-\bar{x}}{\sigma} \quad (19)$$

Où, x est le vecteur de caractéristiques, $\bar{x}$ est la moyenne de x, et σ est son écart-type.

Ensuite, ces vecteurs de caractéristiques normalisés ($y_{Gabor-LDA}, y_{ULBP-LDA}, y_{LPQ-LDA}$) sont concaténés en une seule vecteur z (Fusion au niveaux caractéristiques). Puis une transformation linéaire est appliquée à z afin de le normaliser dans la plage [0 , 1] pour que les caractéristiques soit compatible avec le l'AE.

$$X = \frac{z-z_{min}}{z_{max}-z_{min}} \quad (20)$$

Cette transformation accélère le processus d'encodage et assure une similarité entre l'entrée et la sortie de l'AE, surtout si les fonctions sigmoïdes logistique ou linéaire saturation positive sont utilisées comme fonctions de transfert pour le décodeur. Pour montrer l'effet de ce processus de normalisation, une évaluation a été réalisée sur la base de données YALE comme le montre la Figure II.11.

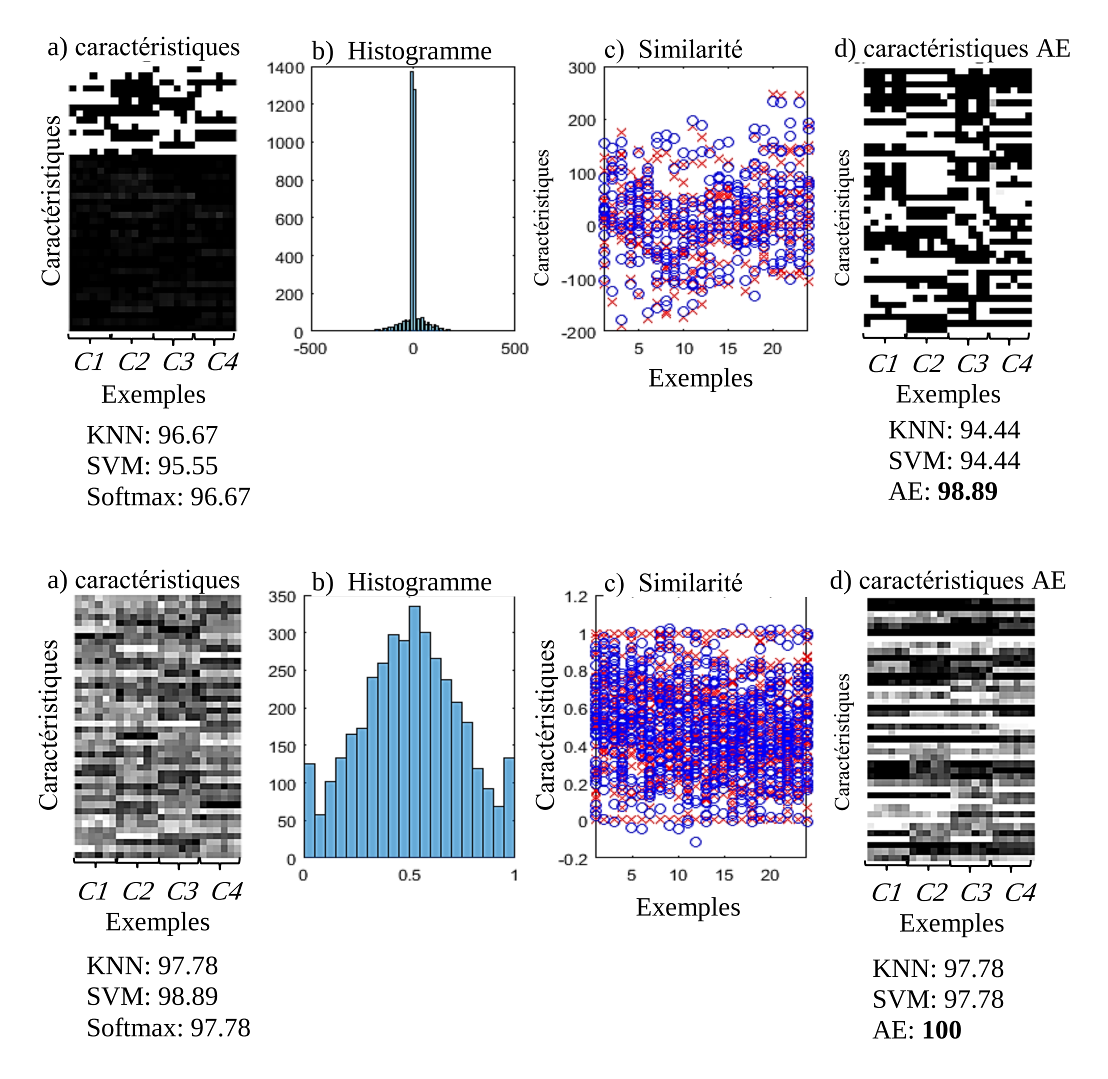


Figure II.11 Le résultat de la normalisation des caractéristiques a montré les quatre premières classes de la base de données de Yale. Sans normalisation en haut, et avec normalisation en bas.

II.4.7 Autoencoder (AE)

Ce travail propose d'entraîner un AE avec deux couches superposées : une couche d'encodeur pour encoder nos combinaisons de caractéristiques x dans les nœuds cachés h et une couche de classificateur softmax pour la classification. La Figure II.12 montre les trois étapes du processus d'entraînement proposé.

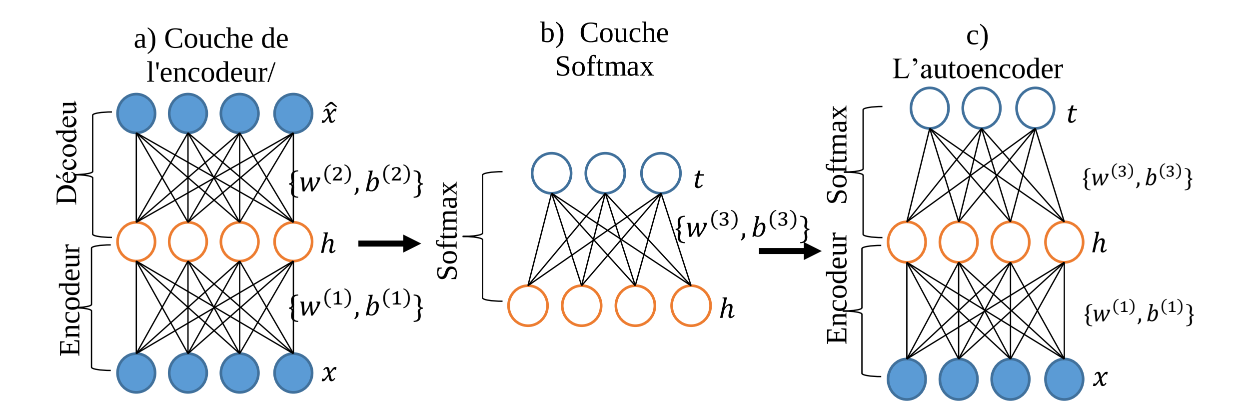


Figure II.12 L'architecture en trois étapes pour l'entraînement l'autoencoders

II.4.8 Phase d'entraînement de l'encodeur

Premièrement, dans la phase d'entraînement de la couche de l'encodeur, nous avons entraîné un modèle AE composé de deux couches : un encodeur et un décodeur sont entraînés pour faire correspondre l'entrée x à sa sortie $\hat{x}$, Figure II.12(a). Un Backpropagation basé sur gradient conjugué avec une fonction de coût basée sur l'erreur quadratique moyenne, la régularisation L_2, et la régularisation de la sparsité, ont été utilisés pour assurer la proximité entre l'entrée et sa valeur reconstruite dans la sortie [100]. Ici, l'encodeur fait correspondre l'entrée x aux nœuds cachés h par la fonction intermédiaire f :

$$f\colon h = f(x) \quad (21)$$

Le décodeur renvoie les nœuds cachés à l'espace d'entrée d'origine x par une autre fonction de correspondance déterministe g :

$$g\colon \hat{x} = g(h) \quad (22)$$

La fonction de coût global utilisée dans la rétro-propagation [101] est :

$$E = MSE + \lambda \times \Omega_{weights} + \beta \times \Omega_{sparsity} \quad (23)$$

Où λ et β sont les coefficients de la régularisation L_2 et de la régularisation de la sparsité, respectivement. Dans ce travail, nous avons posé $\lambda = 10^{-6}$ comme valeur par défaut pour augmenter l'importance du terme de sparsité [102] et prenons β variable dans les résultats de l'expérience. L'erreur quadratique moyenne de la fonction de coût définie par l'équation :

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K}(x_{kn} - \hat{x}_{kn})^2 \quad (24)$$

Où N est le nombre d'observations, et K est le nombre de variables dans les données de sortie, en outre, le terme de régularisation L_2 est :

$$\Omega_{weights} = \frac{1}{2}\sum_{l}^{L}\sum_{n}^{N}\sum_{k}^{K}\left({w_{nk}}^{(l)}\right)^2 \tag{25}$$

Où $*^{(l)}$ est le numéro du modèle l ($l = 1$: le premier codeur, $l = 2$: le premier décodeur), L est le nombre de couches cachées, et $w^{(l)}$ est la $l^{-ième}$ matrice de poids. Enfin, le terme de la sparsité régularisation de l'équation (26) est ajouté pour contrôler la sparsité de la sortie de la couche cachée. La sparsité peut être contrôlée en ajoutant le terme divergence de Kullback-Leibler (K L) [101], qui prend une grande valeur lorsque $\hat{p}_i$ d'un neurone i et sa valeur désirée p ne sont pas proches en valeur. Figure II.13(d) ; cela peut entraîner la suppression de neurones [72]

$$\Omega_{sparsity} = \sum_{i=1}^{D^{(1)}} KL(p||\hat{p}_i) = \sum_{i=1}^{D^{(1)}} p\log\left(\frac{p}{\hat{p}_i}\right) + (1-p)\log\left(\frac{1-p}{1-\hat{p}_i}\right) \tag{26}$$

Où D est le nombre de neurones dans la couche cachée, $\hat{p}_i$ est l'activation moyenne de sortie mesure le $i^{-ième}$ neurone de l'encodeur.

$$\hat{p}_i = \frac{1}{N}\sum_{n=1}^{N} h_{ni} \tag{27}$$

Où h_{ni} est la valeur de sortie du $i^{-ième}$ neurone pour chaque entrée x_n des N images d'entraînement.

$$h_{ni} = \varphi^{(1)}(z_{ni}) = ({w_i}^{(1)}x_n + {b_i}^{(1)}) \tag{28}$$

Où $\varphi^{(1)}$ est la fonction de transfert pour le codeur, ${w_i}^{(1)}$ est la $i^{-ième}$ ligne de la matrice de poids du codeur et b_i est la $i^{-ième}$ entrée du vecteur de biais du codeur. Le paramètre prédéfini (proportion de sparsité) est utilisé comme valeur désirée pour chaque $\hat{p}_i$, sa valeur est comprise dans l'intervalle [0, 1] afin de s'assurer que tous les neurones ne se déclenchent pas aux valeurs 0 ou ne sont pas activés à 1 [72]. Dans ce travail, nous avons fixé $p = 0{,}5$ pour que la sortie moyenne des neurones soit proche du centre (0,5) de manière à conserver nos caractéristiques centrées entre 0 et 1, Figure II.13 (a), (b) et (c), cela garantit que nos résultats sont stables si l'on fait varier le coefficient de régularisation de la sparsité β comme le montre la Figure II.13 (e).

Dans ce travail, nous avons choisi la fonction de transfert linéaire à saturation positive pour le codeur, et la fonction de transfert linéaire $\varphi(z)^{(2)} = z$ pour le décodeur. La fonction du décodeur g devient comme indiqué dans l'équation (29) :

$$g: \hat{x}_n = g(h_n) = \varphi^{(2)}\left(w^{(2)}h_n + b^{(2)}\right) = w^{(2)}h_n + b^{(2)} \tag{29}$$

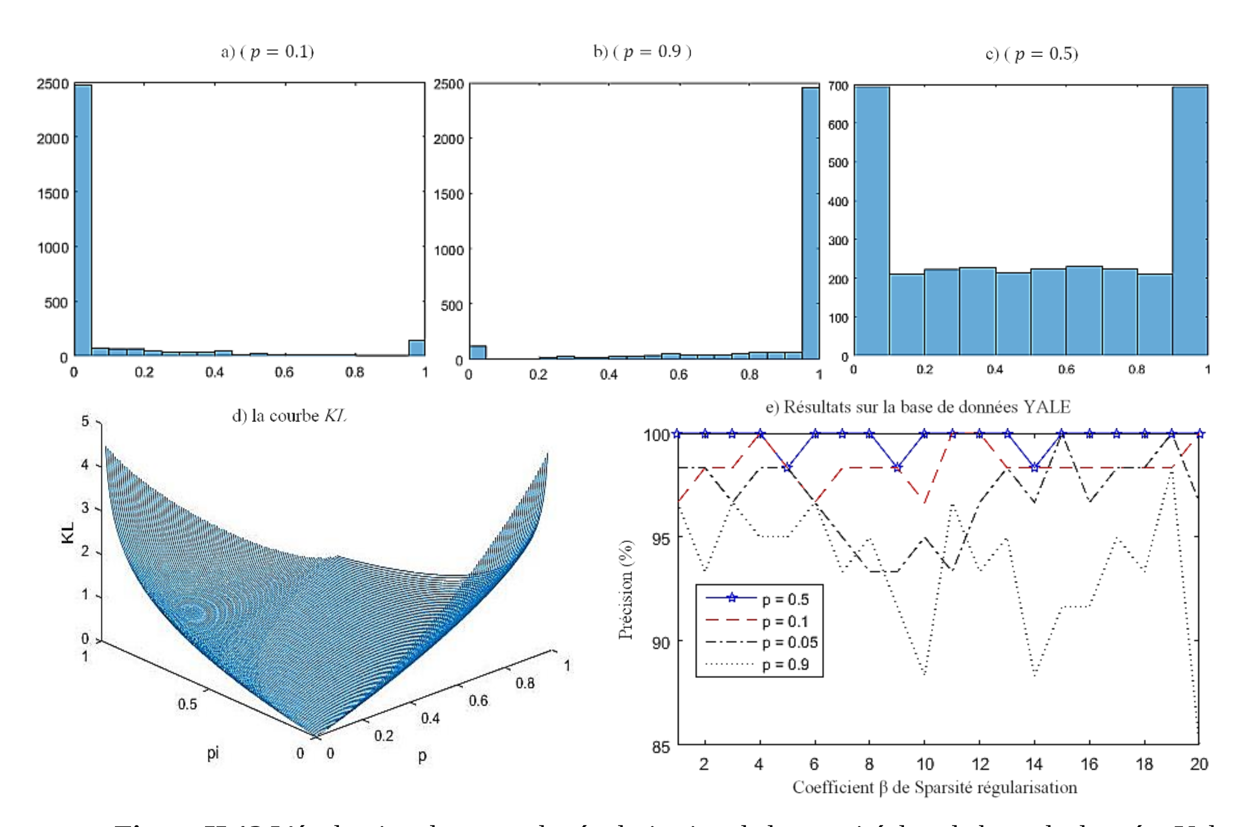


Figure II.13 L'évaluation du terme de régularisation de la sparsité dans la base de données Yale

II.4.9 Phase d'entraînement du classifieur

Dans la deuxième phase, une couche de classification (softmax) est entraînée pour faire correspondre les caractéristiques h à la destination t (classe) Figure II.12(b). La rétro-propagation basée sur le gradient conjugué [100] a été utilisée pour entraîner le classifieur, avec une erreur d'entropie croisée pour calculer l'erreur entre la sortie y et la cible t en utilisant la fonction suivante :

$$E = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K} t_{kn} \ln y_{kn} + (1 - t_{kn}) \ln(1 - y_{kn}) \tag{30}$$

Où N et K sont le nombre d'exemples d'entraînement et le nombre de classes, respectivement

II.4.10 Phase d'entraînement de l'Autoencodeurs en cascade

Enfin, la couche de l'encodeur et la couche du classificateur (softmax) sont combinées dans le même réseau, et en les ré entraînant à nouveau par mappant les caractéristiques x directement a les classes t, afin de générer les nouveaux paramètres du réseau AE $\{w^{(1)}, b^{(1)}\}$ et $\{w^{(3)}, b^{(3)}\}$ [38]. Cette phase garantit que les caractéristiques de la couche cachée h sont plus discriminantes pour les classes t, comme le montre la Figure II.14.

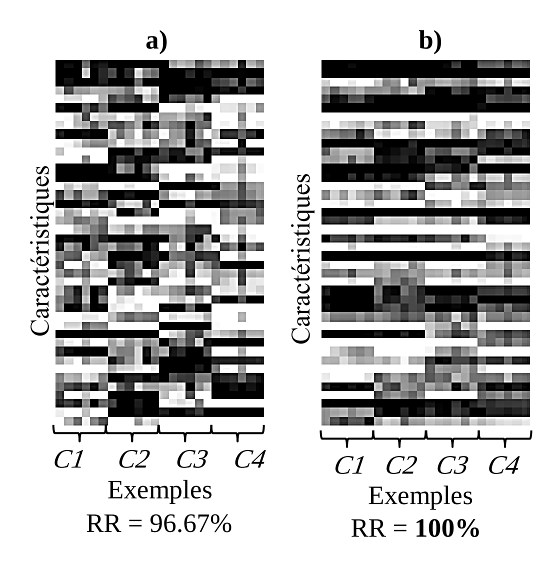


Figure II.14 Visualisation des caractéristiques de la couche cachée h des quatre premières classes de la base de données YALE avec les taux de reconnaissance, (a) avant le réentraînement de l'AE (b) après le réentraînement avec la cible t (classes)

II.5 Conclusion

Dans ce chapitre, nous avons présenté un état de l’art sur la reconnaissance de visage, avec ses application et difficultés, nous avons décrit les différentes étapes qui constituent le système biométrique multi-algorithme proposé pour la reconnaissance de visage tel que le prétraitement, les méthodes d'extraction des caractéristiques (Gabor, LBP, LPQ), la méthode de réduction (LDA), la normalisation, et la classification par l'autoencoder (AE).

Chapitre III

Reconnaissance multimodale:

Fusion avec l'iris

Introduction

L'iris est une structure fine et circulaire de l'œil qui est un organe interne protégé et qui n'est donc pas affecté par les conditions environnementales [103]. Parmi tous les systèmes de reconnaissance biométrique, l'iris est la solution la plus prometteuse en raison de son caractère unique, de sa fiabilité et de sa stabilité au fil du temps. Même des jumeaux génétiquement identiques ont des textures d'iris différentes [104]. Les améliorations technologiques ont conduit à des innovations dans le système de reconnaissance biométrique de l'iris.

En 1936, Frank Burch a adopté le concept d'utilisation des motifs de l'iris pour reconnaître un individu. En 1985, Flom et Safir ont développé le concept selon lequel il n'y a pas deux iris identiques. En 1993, John Daugman a développé un algorithme pour un modèle de reconnaissance automatique de l'iris, dans la même année, le prototype de ce modèle a été testé par l'Agence nucléaire de la défense, grâce aux efforts combinés de Flom, Safir et Daugman qui est terminé avec succès en 1995. Les premiers produits commerciaux sur l'iris ont été mis sur le marché par John Daugman en 1994 [105]. En 2005, le concept de base de la reconnaissance de l'iris, déposé par Flom, a expiré, ensuit le concept de reconnaissance de l'iris basé sur le code de l'iris mis en œuvre par Daugman a expiré en 2011, pour fournir les possibilités de commercialisation à d'autres entreprises pour développer leurs propres algorithmes de reconnaissance de l'iris [106].

En 2011, Ali Alheeti [107] a proposé une technique de reconnaissance de l'iris hybride, qui permet d'identifier la puissance des opérateurs de détection des contours, utilisés pour générer les caractéristiques minimales nécessaires à l'identification d'une iris. Dans cette technique hybride, les transformées en ondelettes discrètes 2D, avec des masques d'ondelettes, comme les masques de Haar et Db2 sont décomposées, suivies par l'application d'opérateurs de détection des contours comme Canny, Prewitt, Roberts et Sobel pour reconnaître les caractéristiques. Rashad et al [108] a proposé LBP avec l'histogramme pour extraire les informations de texture de l'iris, puis pour concevoir un vecteur de caractéristiques, puis introduite dans un classificateur basé sur un réseau neuronal appelé LVQ combiné. Panganiban et al [109] ont implémenté une technique pour acquérir l'image de l'iris en utilisant une caméra vidéo, suivie d'un traitement à l'aide de l'outil d'acquisition d'images MATLAB, ensuite, sur la base de différents coefficients, l'image normalisée a été décomposée à l'aide de Haar et d'ondelettes bi-orthogonales à N niveaux pour extraire les caractéristiques, les résultats ont été

testés sur la base de données CASIA V3. Farouk [110] a proposé un système basé sur la décomposition des ondelettes de Gabor et la correspondance élastique des graphes ; Dans cette méthode, l'iris est segmentée à l'aide de la transformée de Hough circulaire et les ondelettes de Gabor, pour déterminer la similarité et la di-similarité entre deux codes d'iris; L'expérience est testée sur les bases de données CASIA V3 et UBIRIS et UBIRIS.

En 2012 Sathish [111] a proposé un système multi algorithmique, dans lequel l'iris est segmentée en effectuant les étapes suivantes : (1) Initialement, une fonction de filtrage gaussienne puis l'égalisation de l'histogramme est appliquée pour améliorer le contraste de l'image de l'iris. (2) Un détecteur de contours de Canny suivi d'une transformée de Hough circulaire probabiliste est utilisée pour segmenter l'image. (3) L'iris segmenté est ensuite normalisé en utilisant le modèle « rubber sheet » de Daugman, puis les caractéristiques ont été extraites par des filtres de Gabor 2D sur l'image normalisée. (4) Un score de correspondance est obtenu en utilisant la distance de Hamming par le classificateur de réseau neuronal (NN), les résultats ont été testés sur la base de données CASIA. Szewczyk et al [112] utilise des images d'iris acquises dans des conditions non contrôlées, et propose une stratégie pour reconnaître une iris. Dans cette méthode, une signature compressée de 324 bits de largeur est utilisée par rapport à la signature de Daugman de 2048 bits. Le codage de la signature est effectué en utilisant des transformées en ondelettes sur la décomposition de l'image et la technique de binarisation, suivi par le calcul du score pour trouver la correspondance.

En 2013 Zhou et al [113] a proposé une nouvelle technique de comparaison avec codes. Pendant l'étape de segmentation, les étapes suivantes ont été suivies : (i) pour localiser la zone de la pupille, une analyse d'histogramme et un traitement morphologique ont été effectués, (ii) Le contour extérieur a été considéré comme ayant deux fois la taille du contour pupillaire. (iii) Pour détecter et supprimer les paupières supérieures et inférieures, l'opérateur de Canny suivi de l'algorithme d'ajustement de la courbe polynomiale ont été utilisés. Après avoir segmenté l'iris, il a été transformé en un bloc rectangulaire de taille fixe à l'aide d'un opérateur de convolution. Des filtres 1D Log Gabor ont été appliqués pour extraire les informations de texture et ont ensuite été stockés dans une structure arborescente à k-dimensions. À l'aide de cette structure, la correspondance des codes a été effectuée pour trouver la similarité ou la dissimilarité entre deux codes.

En 2014 Rai et al [114] a proposé une technique pour effectuer la correspondance des codes, basée sur la combinaison de deux techniques pour obtenir un meilleur taux de précision. La transformée de Hough circulaire est utilisée pour isoler l'image de l'iris, puis pour trouver la zone de la collerette en zigzag, et enfin pour détecter et supprimer les paupières et les cils en utilisant la technique de détection de la parabole et les filtres médians ajustés. Les ondelettes de Haar et les filtres 1D Log Gabor sont utilisés pour extraire les caractéristiques de la région de la collerette en zigzag de l'iris. Les caractéristiques extraites ont été identifiées à l'aide d'une combinaison de SVM et la distance de Hamming. Les résultats expérimentaux montrent un excellent taux de reconnaissance lorsque les caractéristiques sont extraites de la région spécifique, où des motifs plus complexes sont disponibles. Song et al [115], a proposé une méthode basée sur un modèle de correction d'erreurs sparse, puisque les facteurs de bruit tels que l'occlusion des paupières et des cils et les réflexions spéculaires et pupillaires sont principalement spatiale. Pour rendre ce modèle de correction d'erreurs efficace, un algorithme K-SVD est mis en œuvre, Il est prouvé que les caractéristiques obtenues à l'aide de cet algorithme ont une meilleure représentation. La méthode proposée permet de gagner un temps de calcul considérable et offre un meilleur taux de reconnaissance. Sun et al [116] a fourni un système de classification des images d'iris basé sur les informations de texture avec une technique de représentation appelée Hierarchy Visual Codebook (HVC). HVC est basé sur deux techniques appelées Vocabulary Tree (VT), et Locality-constrained Linear Coding (LLC), pour représenter les textures de l'iris de manière sparse. Le filtre de Gabor et les filtres ordinaux sont utilisés pour extraire les caractéristiques des images d'iris segmentées. Les résultats expérimentaux montrent que cette méthode permet d'obtenir une meilleure classification des images pour la détection de la lividité de l'iris, la classification des races.

La fusion des modalités du visage et de l'iris est une approche biométrique qui a fait l'objet d'une attention croissante au cours de la dernière décennie, probablement en raison de la popularité des modalités individuelles, ainsi que de la connexion naturelle entre elles, qui donne la possibilité d'utiliser un seul capteur. Malgré cette tendance récente, très peu d'études ont été réalisées sur la fusion des données biométriques du visage et de l'iris.

Gawande et al (2014) [88] ont proposé une nouvelle approche en multi-biométrie, , utilisée la programmation linéaire robuste (RLP) pour l'identification à distance. RLP utilise des contraintes incertaines et a été modélisé dans le contexte de la biométrie en concaténant

toutes les caractéristiques (fusion au niveau de caractéristiques) à partir de différentes modalités biométriques. La méthode proposée a été testée sur la base de données CASIA-Iris-Distance et donne une grande précision pour les environnements bruyants, et devrait donc être envisagée pour des recherches plus approfondies sur la reconnaissance à distance où le bruit est un facteur important.

Azom et al (2015) [87] ont proposé un schéma hybride utilisant cinq algorithmes différents d'extraction de caractéristiques : Principal component analysis (PCA), Linear discriminant analysis (LDA), Local binary pattern histogram (LBPH), Sub-pattern principal component analysis (SPCA), et Modular principal component analysis (MPCA) pour le visage et l’iris. Deux fusion au niveau des caractéristiques est effectuée pour chaque modalité (multi-algorithme), ce qui permet d'obtenir deux classificateurs (un pour chaque modalité). Une autre fusion au niveau du score est effectuée pour l'algorithme d'extraction de visage LDA et LBPH pour l'iris (multimodal), créant ainsi le troisième classificateur. Finalement une fusion au niveau de la décision est effectuée sur les résultat du trois classificateurs.

Ryan et all (2016) [26] propose un système qui acquiert simultanément des images de visage et d'iris à l'aide d'un seul capteur, dans le but d'améliorer la précision de la reconnaissance tout en minimisant le coût du capteur et le temps d'acquisition. Le système résultant améliore les taux de reconnaissance par rapport aux taux de reconnaissance observés pour l'autre des biométries isolées.

Ce chapitre présente les différentes étapes de la reconnaissance d'une image d'iris, à savoir l'acquisition, la segmentation, la normalisation, l’extraction de caractéristique, et la classification. Il fournit également le schéma principal de l'application de notre système sur la reconnaissance de l'iris ainsi que le module de fusion proposé avec le visage.

III.1 Propriétés de l'iris

Les principales propriétés de l'iris humain en tant que l'identifiant, que les chercheurs et les développeurs doivent prendre en compte sont les suivantes [106]:

1. L'iris est un organe interne nettement protégé et visible de l'extérieur à une distance pouvant atteindre quelques mètres. Il est protégé derrière la paupière, la cornée et les humeurs aqueuses.

2. Il est hautement texturé avec des formes aléatoires d'une grande complexité et uniques qui génèrent leurs motifs par épigénétique et ont un caractère persistant.

3. L'iris est généralement occultée par les paupières, les cils, les lentilles, les lunettes et les reflets lumineux. Pour certains groupes raciaux, la texture de l'iris est pauvre et partiellement occultée.

4. La texture de l'iris se déforme de manière non élastique lorsque la pupille change de taille.

5. L'iris est petit et une cible mobile à acquérir à distance.

III.2 Les Difficultés dans la reconnaissance de l'iris

La capture d'une image de bonne qualité est l'un des principaux défis de la reconnaissance automatique de l'iris. Le processus d'acquisition de l'image affecte la performance de la localisation de l'iris et a un impact important sur la précision de la reconnaissance [117]. Les facteurs de bruit peuvent être classés en deux catégories : locaux et globaux [118]. La catégorie locale contient les obstructions de l'iris dues aux ; paupières, cils, lunettes, lentilles, cheveux, réflexions de l'éclairage, images spéculaires, hors angle et aux l'iris partielles. La catégorie globale est composée d'images mal focalisées, floues en mouvement, rotation, mal éclairées et hors iris. La Figure III.1 montre les images d'iris qui contiennent ces bruits [119].

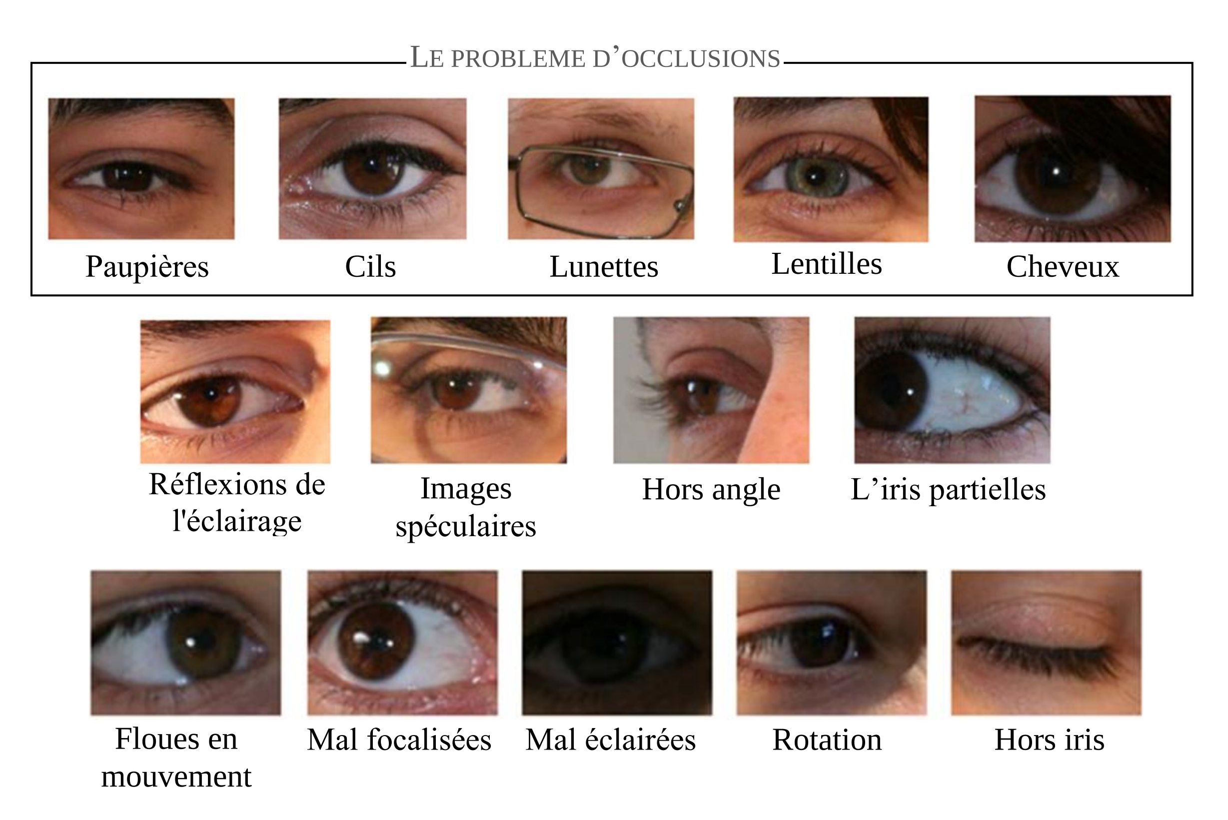


Figure III.1 Exemple des Difficultés dans la reconnaissance de l'iris

1. Le problème d'occlusions

Tous les obstructions par : les paupières, les cils, les lunettes, les lentilles, les cheveux, se produit et contribue à la mauvaise qualité de l'image. Et puis, les performances du système de reconnaissance de l'iris sont dégradées.

2. Éclairage et réflexion spéculaire

L'éclairage et la réflexion spéculaire crée des petites portions de pixels de haute intensité dans l'image de l'iris. En raison d'une source de lumière inappropriée, la réflexion spéculaire se produit dans la zone de l'iris, qui affecte à la segmentation et en suite à la précision de la reconnaissance.

3. Hors angle

En raison d'un mauvais angle d'orientation du capteur, l'image de l'iris est alors hors angle. Dans un tel environnement non coopératif, la longueur de la zone de l'iris est réduite. De plus, la détection des limites devient fastidieuse.

4. Iris partiel

L'image de l'iris partiel est due au fait que le sujet n'a pas focalisé son regard lorsque la caméra a capturé l'image. Dans ce cas, l'image partiellement capturée est considérée comme une situation non idéale. De plus, les performances de la reconnaissance de l'iris diminuent considérablement lorsque l'iris est imprécise.

5. Flou de mouvement

En raison du mouvement de la caméra ou de l'objet pendant l'acquisition de l'image, l'iris devient floue. En résulte, ça produit une fausse identification de la région d'intérêt dans le processus de segmentation.

6. Mauvaise focalisation

Une image de l'iris mal focalisée est capturée avec une perte du focus sur le sujet lors de l'acquisition de l'image de l'iris. Dans ce cas, la précision de la segmentation de l'iris est affectée dans la localisation de la limite interne.

7. Mauvaise luminosité

En raison d'un mauvais éclairage, moins d'informations sont capturées, les limites sont difficiles à identifier et les caractéristiques de la texture peuvent ne pas être clairement mises en évidence. La précision de la reconnaissance s'en trouve réduite.

8. Rotation

La rotation de l'image de l'iris est due au fait que l'axe optique de l'œil n'est pas aligné avec l'axe optique de la caméra. Dans ce cas, les limites de l'iris sont difficiles à identifier dans une forme circulaire.

9. Hors iris

L'image hors iris se produit lorsque le sujet ferme tout simplement les yeux. Cela pose des problèmes de détection des limites de l'iris, car l'image de l'iris n'est pas présente. Pour éviter ce problème, la coopération du sujet est importante.

III.3 L'extraction de la texture de l'iris

Dans tous les system biométrique basé sur la reconnaissance de l'iris, la segmentation et la normalisation joues des rôles très important dans la réussite globale du système. La Figure III.2 ci-dessous présente un schéma fonctionnel de l'extraction de la texture de l'iris.

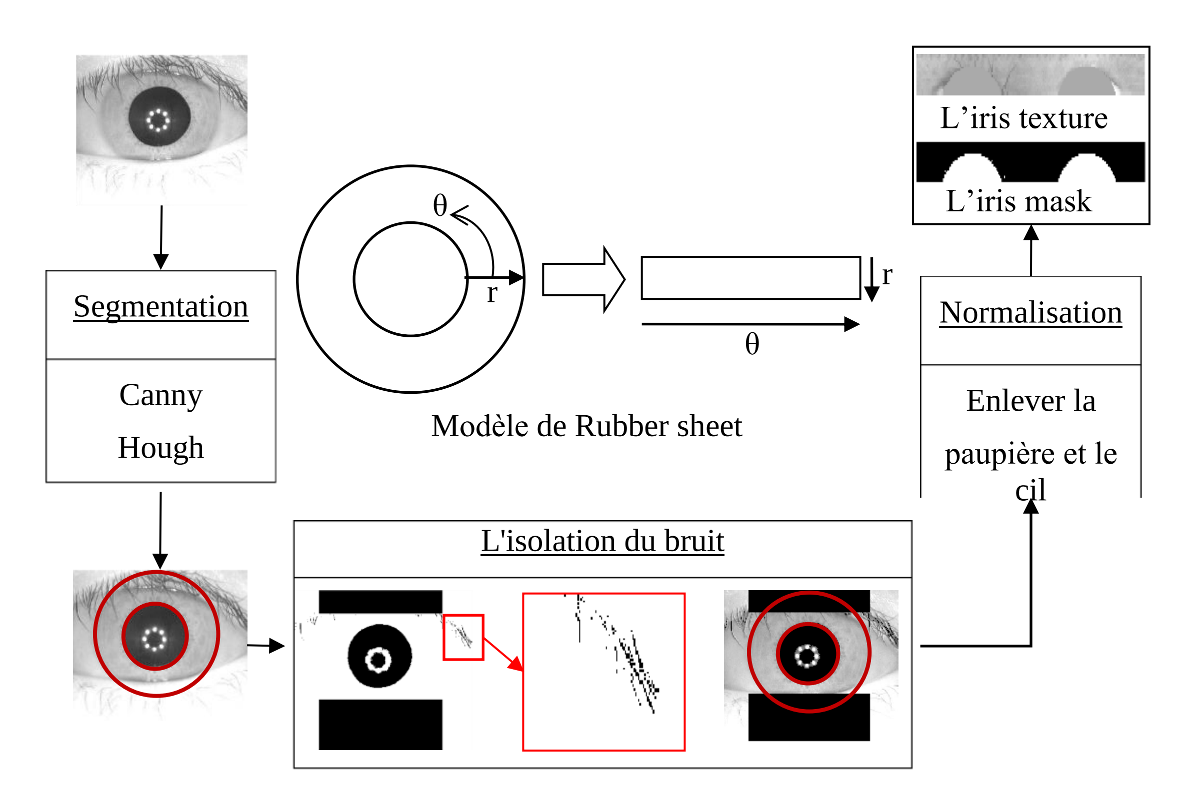


Figure III.2 Schéma de base pour l'extraction de la texture de l'iris

III.3.1 La segmentation

La segmentation de l'iris est effectuée par l'algorithme de Daugman [105]. Il utilise le détecteur de contour de Canny et la transformée de Hough pour déterminer le centre et le radius de l'iris et de la pupille de chaque image d'œil.

La détection de contour Canny a plusieurs étapes. Ce sont respectivement le lissage, la recherche du gradient, la suppression du non-maximum, le seuillage et la détection. Dans la première étape, le bruit de l'image est éliminé en rendant l'image de l'iris floue. Ensuite, après avoir identifié la grande magnitude du gradient de l'image, l'opérateur marque les contours. Après l'achèvement de l'étape précédente, l'opérateur marquera les maximums locaux comme des contours. Avant de déterminer les contours, un seuillage sera utilisé pour identifier les

contours les plus forts et l'étape de détection supprimera les contours faibles qui ne sont pas connectés aux contours forts. Comme l'explique dans [120], l'opérateur Canny est optimal même pour les images bruyantes et à de meilleures performances.

La transformée de Hough est une technique de reconnaissance de formes développée en 1962 par Paul Hough. Cette technique permet de reconnaître des lignes (droite), et des cercles (dans ce cas, les contours de Canny) présente dans une image, et nous donne le centre et le radius de ces cercles. La région de l'iris peut être approximée par deux cercles, un pour la limite iris/sclérotique et un autre à l'intérieur du premier pour la limite iris/pupille [121].

L'isolation du bruit : Les deux paupières (supérieure et inférieure) sont isolées en appliquant la transformée de Radon sur les contours horizontaux [122], [123].

III.3.2 La normalisation

Après avoir extrait les contours et déterminé le centre de l'iris et de la pupille de l'image de l'œil, la forme circulaire de l'iris être transformé sous forme rectangulaire (la texture, Figure III.2) par le modèle Rubber Sheet de Dougman, qui transfère les pixels de la région de l'iris des coordonnées cartésiennes aux coordonnées polaires [105].

III.3.3 La reconnaissance de l’iris : l’application de FAE

Dans les systèmes de reconnaissance de l’iris, en générale les étapes suivantes après l’extraction de la forme rectangulaire de la texture de l’iris, sont la prétraitement (dans certain approche), l’extraction des caractéristiques, et la classification. Dans ce travail nous avons proposé d’utilisé le même système de la reconnaissance du visage « FAE » (les mêmes étapes et les mêmes algorithmes, avec une adaptation de ces paramètres pour l’iris), comme montre la Figure III.3. Cette proposition optimise les ressources d'implémentation (en conséquence le cout d’exploitation) et la vitesse de traitement du système multimodal proposé.

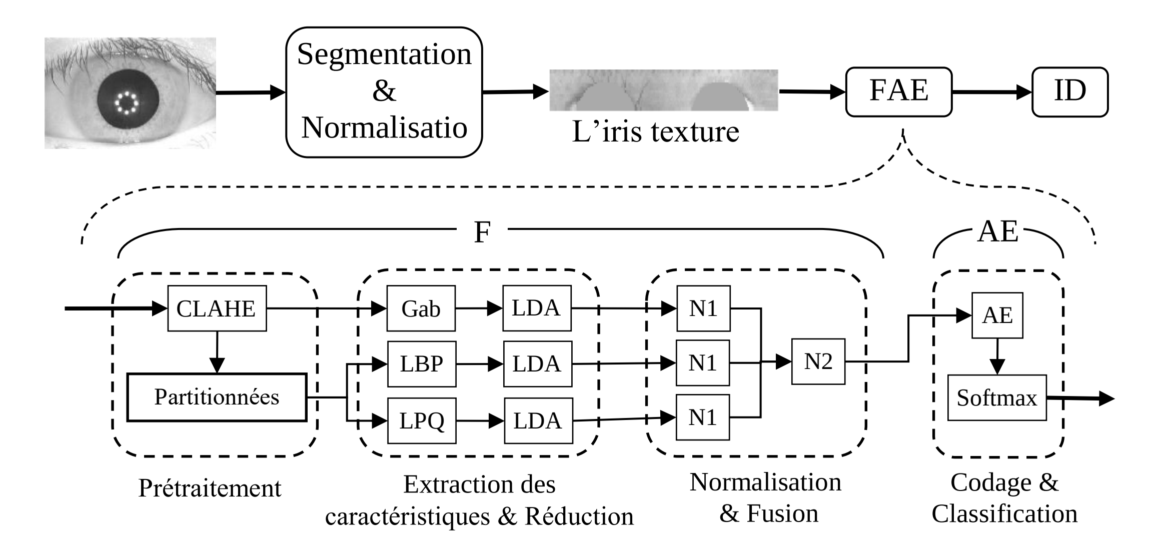


Figure III.3 L'approche proposée pour la reconnaissance de l'iris

III.4 Paramètres adaptation

Parce que les deux modalité (l'iris et le visage) avoir déférant forme, il est obligé d'utilisé les algorithmes avec des paramètres déférant pour augmenter le taux de précision dans la classification (résultats expérimentaux). Les changements effectués sont les suivants :

* Dans le Prétraitement, aucun changement dans l'algorithme CLAHE. Le partitionnement des images est effectué avec une fenêtre de 10x10 sans chevauchement (au lieu de 16x16 dans le cas du visages), c'est parce que, les images de la texture de l'iris sont petites par apport les images des visages, et contient plus de détails.

* Dans l'extraction des caractéristiques, aucun changement dans l'algorithme Gabor. Les algorithmes LBP et LPQ sont effectué sans l'histogramme, en utilise LBP avec R=1 et P=8 (au lieu de R=2, et P=16 dans le cas du visages) c'est parce que, les images de l'iris contiennent une texture complexe discriminant et avec plus de détails (que le visage).

* Le reste des algorithmes sont effectués avec aucun changements (LDA, Normalisation, codage, et la classification).

III.5 Le schéma final : La fusion avec le visage

Le système biométrique peut être divisé en cinq taches principaux : la détection, le prétraitement, l'extraction de caractéristiques, la fusion, et la classification. Dans ce travail, le système biométrique multimodal qui proposé est basé sur le même module proposé d'avance en chapitre 2 (FAE) pour les deux modalités (le visage et l'iris) comme illustré dans la Figure III.4.

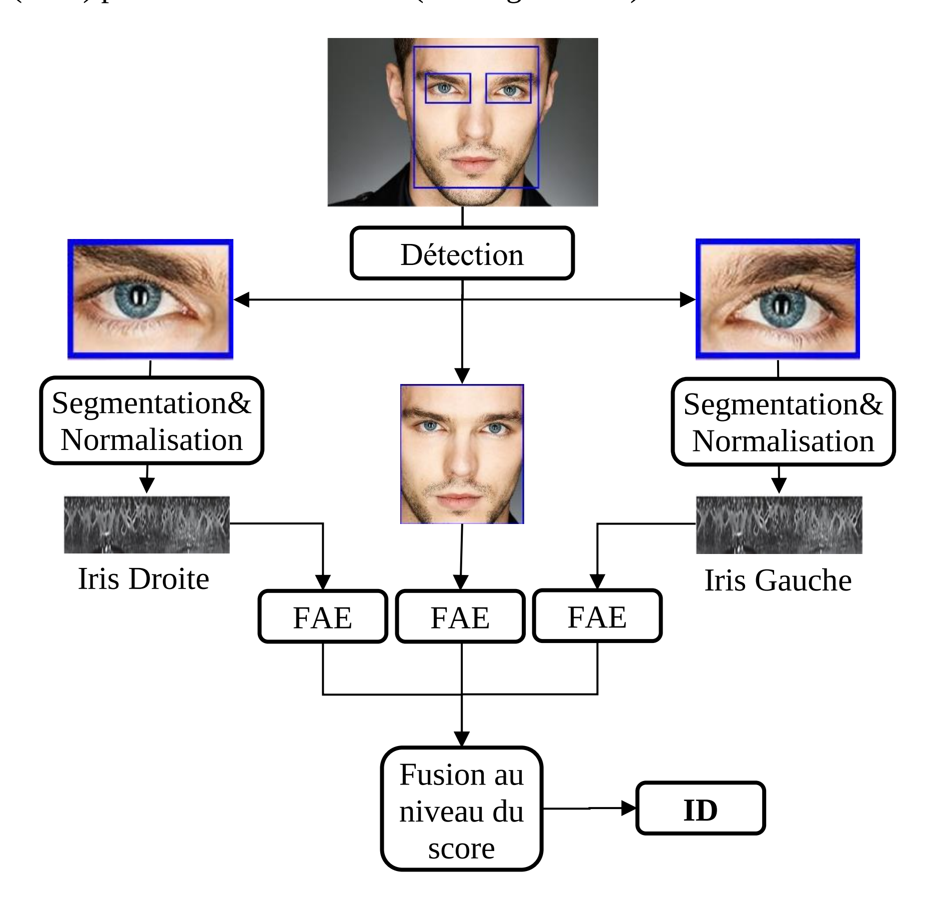


Figure III.4 Le système multimodal proposé

III.5.1 La détection

L'étape de la détection consiste à extraire les régions d'intérêt (ROI) du visage et de l'iris à partir de l'image capturé. Dans ce travail, les images de visage et des yeux gauche et droit sont détectées par le même algorithme de Viola et Jones [124], qui livré (implémentation de OpenCV dans Matlab) avec plusieurs classificateurs pré-entraînés pour la détection des visages frontal, des visages profil, des nez, des yeux et du haut du corps.

III.5.2 La fusion

La stratégie adoptée pour l'intégration dépend du niveau auquel la fusion est effectuée. La fusion au niveau des caractéristiques peut être réalisée en concaténant deux ensembles de caractéristiques compatibles. Des techniques de sélection/réduction des caractéristiques peuvent être utilisées pour traiter le problème de la dimensionnalité.

La fusion au niveau du score de correspondance a été bien étudiée dans la littérature [125], [126]. Des techniques de normalisation robustes et efficaces sont nécessaires pour transformer les scores de plusieurs apparieurs en un domaine commun avant de les consolider [127]. Des stratégies générales pour combiner plusieurs classificateurs ont été proposées dans [128], [129]. Ross et Jain ont montré [130] que la simple règle de la somme est suffisante pour obtenir une amélioration significative de la performance de correspondance d'un système biométrique multimodal. Ils proposent également une technique permettant d'incorporer des poids spécifiques à l'utilisateur pour améliorer encore les performances du système [131]. Les stratégies de fusion au niveau décisionnel comprennent le vote majoritaire [132], la méthode de l'espace de connaissances comportementales [133], le vote pondéré basé sur la théorie de la preuve de Dempster-Shafer [134], les règles ET/OU [135], etc.

Dans notre travaille, le problème comporte à la fois un système multi-algorithmes (c'est-à-dire la fusion entre les algorithme Gabor, LBP, LPQ) un système multi-échantillons (c'est-à-dire les deux iris) et un système multimodal (c'est-à-dire des échantillons d'iris et de visage). Par conséquent, de nombreuses méthodes peuvent être utilisées pour combiner les données biométriques du visage et de l'iris. Dans ce travaille nous propose une fusion au niveau de caractéristiques pour le multi-algorithmes, et une fusion au niveau du score pour le multimodal par la Règle du maximum (max rule) :.

$$S = \max(S_{\text{iris.D}}, S_{\text{iris.G}}, S_{\text{Visage}})$$

III.6 Conclusion

A travers ce chapitre, nous avons présenté un état de l'art sur la reconnaissance de l'iris, et sur la biométrie multimodale visage-iris, les propriétés de l'iris et ses difficultés dans la reconnaissance. Nous avons présenté le système de reconnaissance de l'iris basé sur l'extraction de la texture de l'iris, et la classification par le module proposé de l'apprentissage approfondi (Autoencoder), la deuxième partie de ce chapitre a présenté le schéma final de la fusion.

Chapitre IV

Etudes Expérimentales

IV.1 Introduction

Dans les chapitres précédents, nous avons présenté le système généraliste proposé (FAE), et leurs applications sur la reconnaissance du visage et de l'iris, nous avons aussi parcouru l'implémentation et les paramètres de toutes les méthodes utilisées : le prétraitement, CLAHE, filtrage de Gabor, LBP, LPQ, LDA, l'AE, la normalisation, le codage et la classification.

Dans ce chapitre, nous allons tester, et comparer notre système proposé et ces paramètres et méthodes sur des bases de données biométriques à savoir: le visage et l'iris afin de mettre en évidence leur performances et leur efficacité dans l'identification des individus.

IV.2 Expérimentations sur le visage

Cette section fournit les détails d'implémentation et le paramétrage utilisé pour l'identification de visage, Nous démontrons d'abord la nécessité et l'effet de chaque étape de notre système, où KNN, SVM et Softmax ont été implémentés pour valider la puissance de notre proposition (FAE) pour le codage et la classification des caractéristiques. Dans ces expériences, toutes les images de toutes les bases de données sont converties en niveaux de gris, coupées et redimensionnées à 64 x 64 pixels. La performance de notre système est calculée en utilisant la précision de l'identification calculée par l'équation.

$$\mathrm{Prs} = 100 \times \frac{\mathrm{NC}}{\mathrm{NT}} \qquad (\%)$$

Avec, NC est le nombre des images bien identifiés, NT le nombre totale des images de test.

Trois bases de données accessibles au public, ORL, Yale et AR, ont été utilisées pour évaluer les performances du système proposé. Les tableaux 1, 2 et 3 montrent les résultats expérimentaux, lorsque les méthodes proposées: de prétraitement (CLAHE), de partitionnement (Part), de chevauchement (Chv) et de normalisation (Norm) sont utilisées ou non, dans lesquelles la colonne des caractéristiques : GL, BL, et QL font référence à Gabor, LBP, et LPQ, suivie par LDA, respectivement.

IV.1.1 Résultats expérimentaux sur la base de données de visages ORL

La base de données ORL comprend 400 images de visage, prises sur 40 personnes, chacune ayant 10 images. Les images ont été prises à différents moments, avec des conditions

variables : d'éclairage, d'expressions faciales (yeux ouverts/fermés, souriant/non souriant) et de détails du visage (avec/sans lunettes) (cf. Figure IV.1).



Figure IV.1 Un exemple des visages de la base de données ORL

Pour les expériences, comme utilisé par la plupart des chercheurs, nous avons choisi les quatre premières images pour l'entraînement et les six autres pour le test. Tableau IV.1 montre les résultats obtenus.

Tableau IV.1 Les résultats sur la base de données ORL

Caractéristiques					Précision (%)			
CLAHE	Part	Chv	Caractéristiques	Norm	KNN	SVM	Softmax	**AE**
Non	Non	Non	GL+BL+QL	Non	98.33	96.25	**99.17**	97.91
Non	Non	Non	GL+BL+QL	Oui	97.5	94.58	**98.33**	97.92
Non	16x16	Non	GL+BL+QL	Non	98.33	69.25	99.17	**99.17**
Non	16x16	Non	GL+BL+QL	Oui	100	99.17	99.58	**100**
Oui	16x16	Non	GL+BL+QL	Non	98.33	97.5	98.33	**99.58**
Oui	16x16	Non	GL+BL+QL	Oui	97.5	97.08	98.75	**98.75**
Non	16x16	50%	GL+BL+QL	Non	98.33	96.25	99.17	**99.17**
Non	16x16	50%	GL+BL+QL	Oui	**99.58**	99.17	99.58	99.17
Oui	16x16	50%	GL+BL+QL	Non	98.33	97.5	98.33	**99.58**
Oui	Non	Non	GL	Oui	97.91	96.25	98.33	**98.33**
Oui	16x16	50%	BL	Oui	**97.5**	93.75	97.08	95
Oui	16x16	50%	QL	Oui	**100**	97.92	**100**	98.75
Oui	16x16	50%	GL+BL	Oui	97.5	95.83	98.75	**98.33**
Oui	16x16	50%	GL+QL	Oui	99.58	99.17	99.58	**99.58**
Oui	16x16	50%	BL+QL	Oui	99.17	98.33	99.17	**99.58**
Oui	**16x16**	**50%**	**GL+BL+QL**	**Oui**	**100**	97.92	99.58	**100**

IV.1.2 Résultats expérimentaux sur la base de données Yale

La base de données de visages de Yale contient 165 images en niveaux de gris de 15 personnes, chaque personne ayant 11 images de visages qui comprennent des variations dans les expressions et d'illumination (cf. Figure IV.I.2).



Figure IV.I.2 Un échantillon des visages de la base de données Yale

Pour cette base de données, comme utilisée par la majorité des chercheurs, nous avons choisi les 5 premières images pour l'entraînement et les 6 autres pour le test. Tableau IV.2 montre les résultats obtenus.

Tableau IV.2 Les résultats sur la base de données de Yale

Caractéristiques					Précision (%)			
CLAHE	Part	Chv	Caractéristiques	Norm	KNN	SVM	Softmax	**AE**
Non	Non	Non	GL+BL+QL	Non	**96.67**	95.56	**96.67**	88.89
Non	Non	Non	GL+BL+QL	Oui	93.33	**94.44**	92.3	93.33
Non	16x16	Non	GL+BL+QL	Non	90	83.33	90	**94.44**
Non	16x16	Non	GL+BL+QL	Oui	94.44	94.44	94.44	**95.56**
Oui	16x16	Non	GL+BL+QL	Non	96.67	95.56	96.67	**96.67**
Oui	16x16	Non	GL+BL+QL	Oui	94.44	93.33	94.44	**96.76**
Non	16x16	50%	GL+BL+QL	Non	90	83.33	90	**94.44**
Non	16x16	50%	GL+BL+QL	Oui	95.56	94.44	95.56	**95.56**
Oui	16x16	50%	GL+BL+QL	Non	96.67	95.56	96.67	**98.89**
Oui	Non	Non	GL	Oui	**96.67**	95.56	**96.67**	95.56
Oui	16x16	50%	BL	Oui	**93.33**	**93.33**	**93.33**	88.89
Oui	16x16	50%	QL	Oui	93.33	**95.56**	94.44	93.33
Oui	16x16	50%	GL+BL	Oui	97.78	96.67	97.78	**97.78**
Oui	16x16	50%	GL+QL	Oui	97.78	**98.89**	**98.89**	96.67
Oui	16x16	50%	BL+QL	Oui	94.44	**95.56**	94.44	93.33
Oui	**16x16**	**50%**	**GL+BL+QL**	**Oui**	**97.78**	**98.89**	97.78	**100**

IV.1.3 Résultats expérimentaux sur la base de données AR

La base de données AR se compose de plus de 4000 images frontales de 126 personnes. Nous avons choisi un sous-ensemble contenant 50 sujets féminins et 50 sujets masculins, comme utilisé par la plupart des autres chercheurs. Pour chaque individu, 26 images capturées lors de deux sessions différentes, avec des expressions faciales, des conditions d'éclairage, et des différentes occlusions (cf. Figure IV.I.3).



Figure IV.I.3 Un exemple des visages de la base de données AR

Dans ces expériences, 7 images non occluses sont utilisées comme ensemble d'entraînement et les 6 images occluses restantes avec lunettes et foulards sont utilisées comme ensemble de test. Le tableau IV.3 montre les résultats obtenus.

Tableau IV.3 Les résultats sur la base de données AR session 1

Caractéristiques					Précision (%)			
CLAHE	Part	Chv	Caractéristiques	Norm	KNN	SVM	Softmax	**AE**
Non	Non	Non	GL+BL+QL	Non	83.83	81.5	79.83	**88.83**
Non	Non	Non	GL+BL+QL	Oui	66.5	63.67	**82.83**	52.67
Non	16x16	Non	GL+BL+QL	Non	83.83	81.5	79.83	**97.17**
Non	16x16	Non	GL+BL+QL	Oui	91	85	97.33	**98**
Oui	16x16	Non	GL+BL+QL	Non	96.5	95.17	96.33	**98.83**
Oui	16x16	Non	GL+BL+QL	Oui	98.5	94	**99.33**	98.67
Non	16x16	50%	GL+BL+QL	Non	83.83	81.5	79.83	**98.5**
Non	16x16	50%	GL+BL+QL	Oui	96.5	72.17	98.17	**98.83**
Oui	16x16	50%	GL+BL+QL	Non	96.5	95.17	96.33	**99.33**
Oui	Non	Non	GL	Oui	93.67	89.67	95.5	**95.67**
Oui	16x16	50%	BL	Oui	93.67	90	**98.33**	95.5
Oui	16x16	50%	QL	Oui	**99**	96.83	99.16	98.5
Oui	16x16	50%	GL+BL	Oui	98	94.83	**99.33**	98.83
Oui	16x16	50%	GL+QL	Oui	99	96.83	**99.33**	98.17
Oui	16x16	50%	BL+QL	Oui	98.5	97	99.5	**99.5**
Oui	**16x16**	**50%**	**GL+BL+QL**	**Oui**	98.83	**97.67**	**99.66**	**99.66**

Comme le montre la deuxième partie des tableaux précédents (1, 2 et 3), les résultats obtenus en combinant les trois caractéristiques Gabor + LBP + LPQ sont nettement plus performants que ceux obtenus en utilisant une ou deux caractéristiques séparément. Le prétraitement affecte également les performances de reconnaissance, surtout dans le cas de la base de données Yale, qui présente une grande variation d'illumination. En plus, les tableaux 1, 2 et 3 montrent que l'AE donne de meilleures performances comparées par les autres classificateurs dans la plupart des cas.

Pour expliquer ces résultats, le Tableau IV.4 résume tous les résultats expérimentaux en termes du taux d'amélioration de la classification. Ceci est illustré par (Pre_{avec}), qui est la précision moyenne lors de l'utilisation d'une étape proposée, et (Pre_{sans}) lorsque l'étape proposée n'est pas utilisée, les précisions moyennes ont été calculées statiquement à partir des tableaux précédents (1, 2 et 3).

Tableau IV.4 L'amélioration de la précision par toutes les parties proposées dans Notre approche sur toutes les bases de données (Yale, ORL, et AR)

Étapes proposées (Méthodes)	Pre_{sans}% (Moy /STD)	Pre_{avec}% (Moy /STD)	Gain %
Prétraitement	91.83/9.28	97.29/2.25	+5.46
Normalisation	93.76/6.24	96.02/6.29	+2.26
Part (16x16)	91.12/10.64 (Non Part)	97.06/2.97	+5.94
Chv (50%)	95.20/4.97 (Non Chv)		+1.86
AE	95.00/6.14	96.41/6.87	+1.41
Tous (FAE)	94.88/6.59	**99.89/0.2**	**+5.01**

Comme le montre le Tableau IV.4, on peut dire que l'approche **FAE** proposée améliore les taux de l'identification de 94,88% à 99,89% en utilisant ces bases de données. En outre, lorsque comparés le modèle AE par rapport aux classificateurs KNN, SVM et Softmax, l'AE améliore la performance de 1.41%. En plus, une amélioration des performances de 5.46%, 2.26%, 5.94% et 1.86% est obtenue lorsque : le prétraitement (CLAHE), la normalisation, le partitionnement, et le chevauchement (à 50%) sont utilisés respectivement. Enfin, le tableau 4 montre un gain d'amélioration de +5.01% lorsque tous les caractéristiques (Gabor, LBP, et LPQ) sont utilisées.

IV.1.4 Etude comparative

Cette section analyse les résultats de l'étude comparative de la méthode proposée FAE par rapport à différentes techniques similaires de l'état d'art présentées dans l'introduction. Le Tableau IV.5 présente les résultats obtenus en utilisant des paramètres similaires à l'état d'art, avec les bases de données suivantes :

- **LFWcrop :** constitue la version coupée [136] de la base des données Labeled Faces in the Wild (LFW) [137], qui contient 13233 images de visages collectées sur le Web de 5749 personnes différentes. Les images de cette base de données présentent de riches variations intra-personnelles de pose, d'illumination et d'expression. Elle a été largement étudiée pour la

recherche de la reconnaissance des visages sans contrainte ces dernières années. Dans cette expérience, suivant Liu et al. (2018) [74], un sous-ensemble de 2278 images de personnes qui ont plus de 20 photos, mais moins de 100 photos, a été choisi. N = 90 % d'images choisies aléatoirement pour l'entraînement, et les restants T = 10% de chaque classe pour le test.



Figure IV.I.4 Un exemple des visages de la base de données LFWcrop

- **CMU PIE:** Cette base de données contient 41368 images de 68 personnes, où chacune est prise sous 43 éclairages différents, 4 expressions différentes et 13 poses différentes. Dans cette expérience nous avons pris les 5 poses proches du front, avec toutes les illuminations et expressions (c05, c07, c09, c27 et c29) qui sont fréquemment utilisées [138].



Figure IV.I.5 Un exemple des visages de la base de données CMU PIE

- **AR1 (changement de l'âge):** La base de données AR définie dans la sous-section IV.1.3 , comme décrit dans [52], [69], [139], un sous-ensemble composé de 14 images non occluses de deux saisons a été sélectionné.
- **AR2 (occlusion):** Comme décrit dans [139], 8 échantillons par personne d'images non occluses avec différentes expressions de la base de données AR (section IV.1.3) sont utilisés pour l'apprentissage, tandis que les autres, occluses par l'écharpe et des lunettes, sont utilisés pour les tests.
- **AR3 (illumination et occlusion sans changement dans l'âge):** En utilisant le protocole de [51], nous avons sélectionné 13 images de la saison 1 pour tester notre méthode, parmi les images originales normalisées à 128 × 128, nous avons sélectionné N = 6 échantillons de chaque classe pour l'entraînement, et les autres (T = 7) pour le test.

- **AR4:** Toutes les 26 images sont utilisées à partir des saisons 1 et 2 comme décrit dans [53], [67] , où chaque image est coupée et redimensionnée à 50 × 40 pixels, N est fixé à 9 pour l'entraînement et T = 17 pour les tests.
- **AR5:** Dans cette expérience, un sous-ensemble tel que décrit dans [73], [74], est utilisé pour chaque personne, seules les images avec des changements d'illumination et d'expression ont été utilisées. 7 images de la session 1 sont utilisées pour l'entraînement (N = 7), et les 7 autres images de la session 2 pour les tests (T = 7).
- **Extended YALE B (illumination)** [140] : se compose de 2414 images de visages de 38 individus. Chacun possède environ 64 images proches du front sous différents éclairages.



Figure IV.I.6 Un exemple des visages de la base de données Extended YALE B

Pour évaluer l'efficacité du FAE proposée, différentes valeurs allant de 10 à 20 ont été attribuées aux coefficients de régularisation de la sparsité β, et en sélectionnant aléatoirement N images pour l'apprentissage, et en conservant les images restantes T pour l'évaluation. Les expériences sont répétées 20 fois indépendamment, ce qui génère 220 sous-tests. Les résultats sont ensuite moyennés et les écarts types (STD) ont été calculés comme indiqué dans la colonne 4. Dans le Tableau IV.5, les paramètres et les résultats des travaux relatifs sont présentés dans les colonnes 2 et 3, respectivement. F signifie que l'utilisation des premiers N images pour l'entraînement, tandis que R signifié une sélection aléatoire.

Comme le montre le Tableau IV.5, les taux de reconnaissance (TR) de la FAE proposée sont supérieurs à ceux de toutes les autres techniques. Par exemple, dans le cas de l'illumination et le changement d'expressions (en utilisant les bases de données CMU PIE, Yale étendu et AR1). L'amélioration des performances est significative, qui atteint 99,40 % avec l'utilisation de FAE sur la base Yale étendu, contre 95,99 % et 93,91 % respectivement pour les méthodes GELM et EPCANet. En outre, un TR de 97,31% en utilisant tout la base de données CMU PIE, et de 98,68 % sur le sous-base C29. Par contre un TR de GELM de 93,47 % et 90,60 % pour GELM et LSPDA, respectivement. Pour les distorsions d'occlusion (AR2), les expériences ont été réalisées avec des foulards (TR =98,31%), et un TR= 99,10% lorsque des lunettes sont

incluses. Pour les environnements non contraints sur la base de données LFWcrop, un TR de 90,68% a été atteint par le FAE, par rapport à 88,53% avec EPCANet. Enfin, l'algorithme proposé a atteint les meilleures performances de reconnaissance en utilisant des bases de données avec moins de variation de poses, comme dans le cas de Yale et ORL. Par conséquent, cela démontre clairement que la méthode proposée performe mieux que les méthodes existantes.

Tableau IV.5 La comparaison des performances avec les études relatives.

Auteurs (année)	Base de données et les paramètres	Techniques Évaluation (TR %) et < Longueur des caractéristiques >			**FAE Proposé** TR ± STD, <LongCaract>
Huang et al. (2014) [52]	ORL(112x92), N=F5, T=5	LDP (91.0) <36>	LPP (90.5) <52>	LSPDA (95.5) <26>	**100** (64x64) <117>
	AR1(50x40), N=5R, T=9	LDP (74.8) <150>	LPP (69.7) <170>	LSPDA (86.6) <195>	**99.39**±0.2 <297>
	PIE-C29 (64x64), N=10R, T=14	LDP (78.7) <95>	LPP (78.9) <120>	LSPDA (90.6) <195>	**98.68**±0.33 <201>
Yu et al. (2014) [66]	ORL(56,46), N=F5, T=5	Improved LBP (91.5) </>			**100** <117>
Huang (2010) [50]	ORL(80x80), N=4R, T=6	2DPCA(90.27)<20>	2DLDA(91.63)<20>	2DPCA+2DLDA(92.30)<20>	**99.18**±1.3(64x64) <117>
	Yale(100x100), N=6R, T=5	2DPCA(92.32)<20>	2DLDA(90.51)<20>	2DPCA+2DLDA(94.88)<20>	**98.11**±1.29(64x64) <42>
Lu et al. (2012) [139]	ORL(32x32), N=5R, T=5	DSNPE (96.0) <40>			**98.68**±1.04 <117>
	Yale(32x32), N=F6, T=5	Gabor + DSNPE +SRC (95.1) <20>			**100** <42>
	AR1(33x24), N=5R, T=9	DSNPE (96.1) <300>			**98.04**±0.36 <297>
	AR2 Scarves (50x40)	DSNPE+NN (20.5) <300>	DSNPE+SRC (59.5) <300>		**98.31**±0.49 <297>
	AR2 Sunglass (50x40)	DSNPE+NN (70.0) <300>	DSNPE+SRC (87.0) <300>		**99.10**±0.33 <297>
Yu et al. (2010) [64]	ORL(64x64), N=4R, T=6	GMPTR + NLDA (97.29) <39>			**99.18**±1.3 <117>
	Yale(64x64), N=5R, T=6	GMPTR + NLDA (93.67) <14>			**97.62**±1.64 <42>
Wen et al. (2012) [51]	ORL(112x92), N=4R, T=6	DV-KPCA (93.45) <80>			**99.18**±1.3(64x64) <117>
	Yale(100x80), N=6R, T=5	DV-KPCA (95,13) <40>			**98.11**±1.29(64x64) <42>
	AR3(128x128), N= 6R, T=7	DV-KPCA (91.45) </>			**99.83**± 0.16 <297>
Huang et al. (2015) [67]	ORL(112x92), N=4R, T=6	TWSBF + LDA (95.04) </>			**99.18**±1.3(64x64) <117>
	AR4(50x40), N=9R, T= 17	TWSBF + LDA (94.23) </>			**99.58**±0.18 <297>
Mandal et al.(2009) [49]	ORL(/), N=5R, T=5	Curvelet+ PCA+ LDA (97.7) <60>			**99.68**±0.68(64x64)<117>
	Yale(/), N=6R, T=5	Curvelet+ PCA+ LDA (92.0) <60>			**98.11**±1.29(64x64) <42>
Peng et al. (2015) [73]	ORL(32x32), N=5R, T=5	GELM (96.34) <199>			**98.68**±1.04 <117>
	Yale(32x32), N=5R, T=6	GELM (82.36) <69>			**94.36**±2.32 <42>
	Extended Yale B(32x32),N=20R,T≈44	GELM (95.99) <759>			**99.40**±0.14 <111>
	PIE-All(32x32),N=20R,T≈150	GELM (93.47) <1379>			**97.31**±0.2 <201>
	AR5(60x43), N=7,T=7	GELM(93.85) <300>			**98.80**±1.6 <297>
Fathi et al. (2016) [69]	ORL(92x92), N=5R, T=5	GGZ + HOG (98.0) <90>			**99.68**±0.68(64x64)<117>
	Yale(100x100), N=5R, T=6	GGZ + HOG (97.8) <90>			**98.75**±1.21 <42>
	AR1(120x120),N=7R,T=7R	GGZ + HOG (97.1) <90>			**99.91**±0.13 <297>
Liu et al (2018) [74]	AR5(50,40),N=7,T=7	EPCANet (94.40) </>			**98.64**±0.18 <297>
	Extended Yale B(32x32),N=20,T≈44	EPCANet (93.91) </>			**99.40**±0.14 <111>
	LFWcrop(64x64) ,N=90%, T=10%	EPCANet (88.53) </>			**90.68**±2.45 <183>

IV.2 Les résultats sur l'iris

Pour démontrer la nécessité et l'effet de chaque étape de notre système qui comporte les méthodes de CLAHE, Gabor, LBP, et LPQ. Les performances du système d'identification de l'iris proposé sont évaluées sur la base de données d'iris CASIA-Iris-Interval. Les résultats sont montrés dans le tableau IV.6.

- **CASIA-Iris-Interval-v4**

La base des données CASIA-iris-interval-v4.0 (CASIA-4i) est un sous-ensemble de la base de données CASIA-v4.0. Les images de l'iris de CASIA-Iris-Interval ont été capturées à l'aide de caméra d'iris de proximité, qui peut capturer des images d'iris très claires de 320 × 280 pixels (voir Figure IV.I.7). CASIA-Iris-Interval est bien adapté à l'étude des caractéristiques détaillées de la texture des images de l'iris.

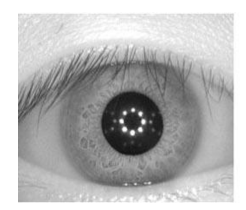
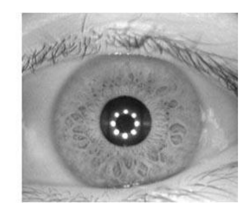



Figure IV.I.7 Exemples d'images d'iris dans CASIA-Iris-Interval.

La base de données complète contient un nombre total de 2639 images (1332 image de l'iris gauche, et 1307 image de l'iris droite) de 249 sujets.

Tableau I.6 Les résultats sur la base de données CASIA-Iris-Interval

	Méthodes	Iris Droite	Iris Gauche	Augmentation Iris Gauche + Droite	Fusion Iris Gauche & Droite
Avec CLAHE	LPQ	88.89	83.16	88.65	98.89
Avec CLAHE	LBP	91.67	88.95	89.46	99.44
Avec CLAHE	LPQ + LBP	96.11	94.74	94.59	99.44
Avec CLAHE	Gabor	98.33	98.33	95.95	**100**
Avec CLAHE	Gabor + LPQ	98.33	98.95	97.84	**100**
Avec CLAHE	Gabor + LBP	98.33	98.95	98.11	**100**
Avec CLAHE	**FAE** (Gabor + LPQ + LBP)	**98.89**	**99.47**	**98.65**	**100**
Gabor + LPQ + LBP | Sans CLAHE		97.22	96.84	97.57	99.44

D'après le Tableau I.6, les meilleures performances d'identification sont celles du système proposé entièrement FAE (Gabor, LBP et LPQ avec le prétraitement CLAHE) avec 98,89% pour l'iris droit, et 99,47% pour l'iris gauche, et 98,65% lorsque nous utilisons une combinaison d'iris gauche et droit pour augmenter la base de données, et nous obtenons **100%** dans le cas de la fusion entre l'iris gauche et l'iris droite. La dernière ligne de la table démontre

que le prétraitement par la méthode CLAHE augmente la précision. D'autre part les deux dernières colonnes de la table (l'augmentation et la fusion de l'iris gauche avec l'iris droite) montrent que les deux iris (gauche et droite) possèdent des caractéristiques totalement différentes, et que l'on peut utiliser les deux dans un système multimodal.

IV.3 Résultats de la fusion visage-iris

Un système biométrique multimodal intégrant les caractéristiques du visage et de l'iris est proposé. Tout d'abord, les caractéristiques du visage et de l'iris sont extraites et classées séparément par la méthode proposé (FAE). Pour prendre la décision multimodale, nous prenons le maximum score (probabilité) parmi les scores des trois modalités (le visage, l'iris droit, et l'iris gauche).

$$S = \max(S_{\text{iris.D}}, S_{\text{iris.G}}, S_{\text{Visage}})$$

IV.3.1 Fusion virtuelle

À cause du manque des bases de données multimodal pour la fusion visage-iris, et dans le but de démontrer que la fusion de visage avec l'iris augment la précision de la reconnaissance, une fusion virtuelle est créée entre la base d'iris (CASIA-Iris-Interval) et les bases de visage qui ont fourni une faible précision (Tableau IV.5). Les expériences de cette section sont réalisées sur la base de donnée CASIA-Iris-Interval, et les bases de données de visage: ORL 32×32, ORL 16×16, Yale 32×32, LFW_Crop 64×64, et CMU PIE 32×32 pour construire une base de données biométrique multimodale expérimentale avec laquelle nous validons l'approche proposée et évaluons la performance de la biométrie multimodale. Le Tableau IV.7 présente les résultats obtenus.

Comme le montre le Tableau IV.7 Les résultats expérimentaux révèlent que l'identification biométrique multimodale est beaucoup plus fiable et précise que l'approche biométrique unique.

Tableau IV.7 Les résultats de la fusion sur des bases de données multimodal virtuelle

	Visage	Iris Gauche	Iris Droite	Iris G & Iris D	Visage & Iris G	Visage & Iris D	Visage &Iris G &D
ORL 32×32	96.67	99.47	98.89	**100**	**100**	**100**	**100**
ORL 16×16	88.89	99.47	98.89	**100**	**100**	**100**	**100**
Yale 32×32	94.67	98.67	94.67	**100**	**100**	**100**	**100**
LFW_Crop 64×64	82.78	98.89	97.22	**100**	**100**	99.44	**100**
CMU PIE 32×32	91.11	98.89	97.22	**100**	**100**	99.44	**100**

IV.3.2 Fusion multimodale

Afin d'évaluer l'applicabilité de la méthode proposée dans cette étude dans des environnements réels multimodal, les expériences de cette section sont réalisées sur une base de données multimodal (visage-iris) publiquement disponible appelée CASIA-Iris-Distance.

- **CASIA-Iris-Distance-v4**

Les images de cette base de donnée ont été capturées par une caméra haute résolution (2352×1728×8b) de sorte que l'iris et le visage sont disponibles dans la région de l'image avec des caractéristiques faciales détaillées, appropriées pour la fusion d'informations biométriques multimodales. Quelques exemples d'images de cette base de données sont présentés à la Figure IV.I.8.



Figure IV.I.8 Un exemple des images dans CASIA-Iris-Distance

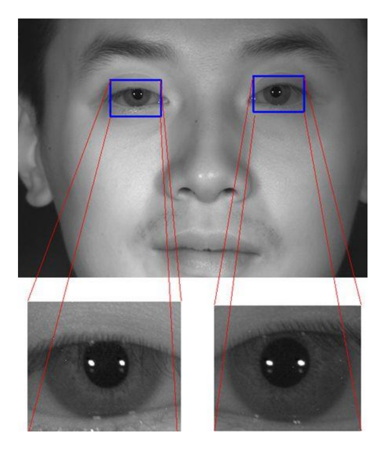


Figure IV.I.9 Un exemple d'iris dans une image de CASIA-Iris-Distance

La base de données complète contient un nombre total de 2567 images de 142 sujets et les images ont été acquises à une distance de ~3 m de la caméra. La taille de la texture des iris extraits dans ce travail est de 240×20. La disponibilité des différentes variations de la base de données CASIA-Iris-Distance est résumée dans le Tableau IV.8 .

Tableau IV.8 Disponibilité de différentes variations pour le visage et les iris dans la base de données CASIA-Iris-Distance

Visage		Iris	
Variations de pose	√	Occlusion - Cils	√
Les expressions faciales	√	Occlusion - Paupière	√
Occlusion - Lunettes	√	Occlusion - Lunettes	√
Occlusion - Moustache	√	Différents facteurs de bruit (reflets, contraste, luminosité, angle mort, rotation, problèmes de flou et de mise au point).	√
Images à distance	√	Images à distance	√

Dans ce travail, nous avons extrait les deux iris de chaque personne à partir de l'image du visage correspondante pour fusionner les modalités du visage et de l'iris. La Figure IV.I.9 montre un exemple de la qualité de l'iris dans une image de CASIA-Iris-Distance.

Tableau IV.9 Les résultats de la fusion sur CASIA-Iris-Distance

Méthodes	Visage	Iris Gauche	Iris Droite	Iris G & Iris D	Visage & Iris G	Visage & Iris D	Visage & Iris G &D
LBP	86.40	56.80	86.80	89.60	94.80	97.60	98.80
LPQ	88.80	87.20	86.80	90	98.40	98.00	98.80
LPQ+LBP	89.20	85.20	86.80	92.00	98.40	97.60	98.80
Gabor	95.20	91.20	86.80	91.60	99.20	99.20	99.20
LPQ+Gabor	94.80	91.20	86.80	92.40	**99.60**	99.20	**99.60**
LBP+Gabor	95.60	90.40	86.80	92.40	**99.60**	**99.60**	**99.60**
Tout (FAE)	94.00	90.80	88.40	92.40	**99.60**	98.80	**99.60**

Le Tableau IV.9 présente les résultats expérimentaux qu'on peut analyser horizontalement et verticalement. Verticalement, il est bien montré (avec le Tableau I.6) que la fusion entre les trois méthodes d'extraction des caractéristiques LPQ + LBP + Gabor, donne les meilleures résultats de précision par rapport la fusion entre deux méthodes, ou d'utilisé un seul méthode. Horizontalement, il est clair que la fusion entre le visage et les deux iris (gauche et droite) donne la meilleure performance dans tous les cas.

IV.4 Conclusion

Dans ce chapitre, nous avons présenté les études expérimentales, comparé et testé les performances de notre système proposé sur des bases de données des visage, iris et aussi sur la fusion multimodale visage-iris. Des comparaisons avec de récents et similaires recherches sont également présentés dans ce chapitre. Les résultats obtenus montrent de manière générale, que

le système proposé avec toutes les méthodes incluses permet de mieux différencier et identifier les individus.

CONCLUSION GENERALE

Ce travail de thèse propose une approche multimodale de la reconnaissance du visage et de l’iris, utilisant une combinaison des caractéristiques améliorée par l'apprentissage approfondi, en utilisant des Autoencodeurs comme méthodologie d'apprentissage approfondi modifiée. L'approche consiste à combiner les caractéristiques locales et globales extraites à l'aide de trois techniques d'extraction de caractéristiques en utilisant une fusion au niveau des caractéristiques. L'entraînement de l'Autoencodeur se fait à l'aide de deux couches, la première fonctionnant comme un codeur et la deuxième comme un classifieur. Ce procédé permet à la technique de mapper l'entrée en utilisant une approche déterministe pour générer les meilleures caractéristiques, et donc de traiter les distorsions du visage et de l’iris d’une manière adéquate. Des expériences détaillées ont été réalisées pour valider l'approche et démontrer son potentiel, comme le montre l'amélioration considérable du taux de reconnaissance (TR) obtenus en utilisant diverses bases de données unimodal et multimodal (visage/iris), et aussi par comparaison avec la communauté des chercheurs. Il a été démontré que la méthode proposée est robuste aux variations d'occlusion, de pose et d'illumination. Les résultats expérimentaux obtenus à l'aide de plusieurs bases de données ont démontré l'efficacité de l'approche proposée qui utilise uniquement les images frontales sans informations supplémentaires provenant des images non contraintes.

En effet, les systèmes combinant plusieurs informations issues de la même biométrie permettent d'améliorer les performances de reconnaissance en réduisant l'effet de la variabilité intra-classe. Cependant, ils ne permettent pas de traiter efficacement tous les problèmes des systèmes monomodaux (par exemple les grandes variations de pose). C'est pour cette raison que les systèmes multi-biométries ont reçu beaucoup d'attention de la part des chercheurs.

Dans un travail plus récent, nous allons créer une architecture CNN pour la grande base de données de la reconnaissance des visages et de l'iris. Dans le cadre d'un travail futur, nous allons étudier l'utilisation de l'alignement des visages pour rendre le système invariant aux grandes variations de pose, nous allons également étudier une sélection de caractéristiques en utilisant différentes méthodes qui sont moins invariantes aux changement de l'âge.

Productions Scientifiques

Publications

Khider, A., Djemili, R.,Bouridane, A. and Jiang, R. (2023) 'Improving face recognition using deep autoencoders and feature fusion', Int. J. Biometrics, Vol. 15, No. 1, pp.40–58. DOI: 10.1504/IJBM.2022.10043147

Conférences

1) A. Benzaoui, A. Khider and A. Boukrouche, "Ear description and recognition using ELBP and wavelets," 2015 International Conference on Applied Research in Computer Science and Engineering (ICAR), Lebanon-Beirut, 2015, pp. 1-6, doi: 10.1109/ARCSE.2015.7338146.

2) Youbi, Zineb, Ali Khider, Larbi Boubchir, Abdelhani Boukrouche, and Badreddine Griouz. "Novel Approach of Face Identification Based on Multi-scale Local Binary Pattern." In 2018 International Conference on Signal, Image, Vision and their Applications (SIVA), pp. 1-4. IEEE, 2018.

3) Ali Khider, Zineb Youbi, Amir Benzaoui, Badreddine griouz, "Face Recognition From Low-Quality Images Based On Deep Learning Autoencoder Model", 29th International Biometric Conference (IBC2018), Barcelona, Spain

4) B Griouz, K Ali, B Abdelhani, Z Youbi 'New approach for finger knuckle-print recognition' XXIX International Biometric Conference (IBC2018), 2018, (Barcelona-Spain)

5) Ali, Khider & Badreddine, Griouz & Abdelhani, Boukrouche & Doghmane, Hakim. (2017). Face recognition based Deep Learning Supervised Autoencoder model. The 5th edition of PhD students day (Oran 2017)

6) Griouz Badreddine, Khider Ali, Bourouba Hocine, Zineb Youbi. 'An efficient algorithme for finger vein recognition'. The 5th edition of PhD students day (Oran 2017)

Références Bibliographiques

[1] H. Imtiaz and S. A. Fattah, 'A curvelet domain face recognition scheme based on local dominant feature extraction', *International Scholarly Research Notices*, vol. 2012, 2012.

[2] A. K. Jain, A. Ross, and S. Prabhakar, 'An introduction to biometric recognition', *IEEE Transactions on circuits and systems for video technology*, vol. 14, no. 1, pp. 4–20, 2004.

[3] A. A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of multibiometrics*, vol. 6. Springer Science & Business Media, 2006.

[4] A. K. Jain, A. A. Ross, and K. Nandakumar, *Introduction to biometrics*. Springer Science & Business Media, 2011.

[5] A. G. Gale and S. S. Salankar, 'Evolution of performance analysis of iris recognition system by using hybrid methods of feature extraction and matching by hybrid classifier for iris recognition system', in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, IEEE, 2016, pp. 3259–3263.

[6] J. Z. Liang and E. Power, 'Iris recognition based on block theory and self-adaptive feature selection', *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8, no. 2, pp. 115–126, 2015.

[7] C. Li, W. Zhou, and S. Yuan, 'Iris recognition based on a novel variation of local binary pattern', *the visual computer*, vol. 31, no. 10, pp. 1419–1429, 2015.

[8] S. Umer, B. C. Dhara, and B. Chanda, 'An iris recognition system based on analysis of textural edgeness descriptors', *IETE Technical Review*, vol. 35, no. 2, pp. 145–156, 2018.

[9] S. Umer, B. C. Dhara, and B. Chanda, 'A novel cancelable iris recognition system based on feature learning techniques', *Information Sciences*, vol. 406, pp. 102–118, 2017.

[10] A. BENZAOUI, 'Identification Biométrique par Descripteurs de Texture Locaux: Application au Visage & Oreille', Thèse de Doctorat, Université 8 Mai 1945 Guelma, 2015.

[11] N. Morizet, 'Reconnaissance biométrique par fusion multimodale du visage et de l'iris', PhD Thesis, Télécom ParisTech, 2009.

[12] R. BELGUECHI, 'Contribution a la reconnaissance d'empreintes digitales par une approche hybride', PhD Thesis, ESI, 2006.

[13] É. Freyssinet and G. Desgens-Pasanau, 'L'identité à l'ère numérique'. Dalloz, 2009.

[14] A. K. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007.

[15] A. Y. J. Nakanishi and B. J. Western, 'Advancing the state-of-the-art in transportation security identification and verification technologies: Biometric and multibiometric systems', in *2007 IEEE Intelligent Transportation Systems Conference*, IEEE, 2007, pp. 1004–1009.

[16] J. L. Wayman, 'Fundamentals of biometric authentication technologies', *International Journal of Image and Graphics*, vol. 1, no. 01, pp. 93–113, 2001.

[17] A. El-Sayed, 'Multi-biometric systems: a state of the art survey and research directions', *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 6, 2015.

[18] Y. Chen, S. C. Dass, and A. K. Jain, 'Fingerprint quality indices for predicting authentication performance', in *International conference on audio-and video-based biometric person authentication*, Springer, 2005, pp. 160–170.

[19] A. S. Nassar, 'A Hybrid Multibiometric System for Personal Identification Based on Face and Iris Traits. The Development of an automated computer system for the identification of humans by integrating facial and iris features using Localization, Feature Extraction, Handcrafted and Deep learning Techniques.', 2018.

[20] F. Matta, 'Video person recognition strategies using head motion and facial appearance', *University of nice sophia-antipolis*, 2008.

[21] K. Nandakumar, 'Multibiometric systems: Fusion strategies and template security', MICHIGAN STATE UNIV EAST LANSING DEPT OF COMPUTER SCIENCE/ENGINEERING, 2008.

[22] Z. Wei, X. Qiu, Z. Sun, and T. Tan, 'Counterfeit iris detection based on texture analysis', in *2008 19th International Conference on Pattern Recognition*, IEEE, 2008, pp. 1–4.

[23] L. Hong and A. Jain, 'Integrating faces and fingerprints for personal identification', *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 12, pp. 1295–1307, 1998.

[24] A. Lumini and L. Nanni, 'Overview of the combination of biometric matchers', *Information Fusion*, vol. 33, pp. 71–85, 2017.

[25] R. Connaughton, *Fusion of face and iris biometrics using a stand-off video sensor*. University of Notre Dame, 2012.

[26] R. Connaughton, K. W. Bowyer, and P. J. Flynn, 'Fusion of face and iris biometrics', in *Handbook of iris recognition*, Springer, 2016, pp. 397–415.

[27] I. H. Sarker, 'Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions', *SN Computer Science*, vol. 2, no. 6, p. 420, 2021.

[28] J. Han, J. Pei, and H. Tong, *Data mining: concepts and techniques*. Morgan kaufmann, 2022.

[29] I. H. Sarker, M. H. Furhad, and R. Nowrozy, 'Ai-driven cybersecurity: an overview, security intelligence modeling and research directions', *SN Computer Science*, vol. 2, pp. 1–18, 2021.

[30] F. Pedregosa *et al.*, 'Scikit-learn: Machine learning in Python', *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[31] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, 'Gradient-based learning applied to document recognition', *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[32] A. Géron, 'Hands-on machine learning with scikit-learn and tensorflow: Concepts', *Tools, and Techniques to build intelligent systems*, 2017.

[33] K. He, X. Zhang, S. Ren, and J. Sun, 'Spatial pyramid pooling in deep convolutional networks for visual recognition', *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, 'Imagenet classification with deep convolutional neural networks', *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[35] K. He, X. Zhang, S. Ren, and J. Sun, 'Deep residual learning for image recognition', in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[36] B. Marlin, K. Swersky, B. Chen, and N. Freitas, 'Inductive principles for restricted Boltzmann machine learning', in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, JMLR Workshop and Conference Proceedings, 2010, pp. 509–516.

[37] G. E. Hinton, S. Osindero, and Y.-W. Teh, 'A fast learning algorithm for deep belief nets', *Neural computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[38] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, 'Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion.', *Journal of machine learning research*, vol. 11, no. 12, 2010.

[39] G. Zhang, Y. Liu, and X. Jin, 'A survey of autoencoder-based recommender systems', *Frontiers of Computer Science*, vol. 14, pp. 430–450, 2020.

[40] A. Makhzani and B. Frey, 'K-sparse autoencoders', *arXiv preprint arXiv:1312.5663*, 2013.

[41] K. W. Bowyer, 'Face recognition technology: security versus privacy', *IEEE Technology and society magazine*, vol. 23, no. 1, pp. 9–19, 2004.

[42] M. De-la-Torre, E. Granger, P. V. Radtke, R. Sabourin, and D. O. Gorodnichy, 'Partially-supervised learning from facial trajectories for face recognition in video surveillance', *Information fusion*, vol. 24, pp. 31–53, 2015.

[43] G. Wang, F. Zheng, C. Shi, J.-H. Xue, C. Liu, and L. He, 'Embedding metric learning into set-based face recognition for video surveillance', *Neurocomputing*, vol. 151, pp. 1500–1506, 2015.

[44] M. A. Muqeet and R. S. Holambe, 'Local appearance-based face recognition using adaptive directional wavelet transform', *Journal of King Saud University-Computer and Information Sciences*, vol. 31, no. 2, pp. 161–174, 2019.

[45] M. Turk and A. Pentland, 'Eigenfaces for recognition', *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[46] B. Schölkopf, A. Smola, and K.-R. Müller, 'Kernel principal component analysis', in *International conference on artificial neural networks*, Springer, 1997, pp. 583–588.

[47] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, 'Eigenfaces vs. fisherfaces: Recognition using class specific linear projection', *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[48] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, 'Face recognition using kernel direct discriminant analysis algorithms', *IEEE transactions on Neural Networks*, vol. 14, no. 1, pp. 117–126, 2003.

[49] T. Mandal, Q. J. Wu, and Y. Yuan, 'Curvelet based face recognition via dimension reduction', *Signal Processing*, vol. 89, no. 12, pp. 2345–2353, 2009.

[50] G. Huang, 'Fusion (2D) 2PCALDA: a new method for face recognition', *Applied Mathematics and Computation*, vol. 216, no. 11, pp. 3195–3199, 2010.

[51] Y. Wen, L. He, and P. Shi, 'Face recognition using difference vector plus KPCA', *Digital Signal Processing*, vol. 22, no. 1, pp. 140–146, 2012.

[52] P. Huang, C. Chen, Z. Tang, and Z. Yang, 'Feature extraction using local structure preserving discriminant analysis', *Neurocomputing*, vol. 140, pp. 104–113, 2014.

[53] Y. Huang and Y. Guan, 'On the linear discriminant analysis for large number of classes', *Engineering Applications of Artificial Intelligence*, vol. 43, pp. 15–26, 2015.

[54] L. Dora, S. Agrawal, R. Panda, and A. Abraham, 'An evolutionary single Gabor kernel based filter approach to face recognition', *Engineering Applications of Artificial Intelligence*, vol. 62, pp. 286–301, 2017.

[55] S. Feng, 'Robust face recognition under varying illumination and occlusion via single layer networks', in *Chinese Conference on Biometric Recognition*, Springer, 2016, pp. 93–101.

[56] T. Ojala, M. Pietikainen, and T. Maenpaa, 'Multiresolution gray-scale and rotation invariant texture classification with local binary patterns', *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[57] D. G. Lowe, 'Distinctive image features from scale-invariant keypoints', *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[58] V. Štruc and N. Pavešić, 'The complete gabor-fisher classifier for robust face recognition', *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 1–26, 2010.

[59] V. Ojansivu and J. Heikkilä, 'Blur insensitive texture classification using local phase quantization', in *International conference on image and signal processing*, Springer, 2008, pp. 236–243.

[60] J. Kannala and E. Rahtu, 'Bsif: Binarized statistical image features', in *Proceedings of the 21st international conference on pattern recognition (ICPR2012)*, IEEE, 2012, pp. 1363–1366.

[61] N. Dalal and B. Triggs, 'Histograms of oriented gradients for human detection', in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Ieee, 2005, pp. 886–893.

[62] M. S. Aslan, Z. Hailat, T. K. Alafif, and X.-W. Chen, 'Multi-channel multi-model feature learning for face recognition', *Pattern Recognition Letters*, vol. 85, pp. 79–83, 2017.

[63] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, 'Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition', in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, IEEE, 2005, pp. 786–791.

[64] L. Yu, Z. He, and Q. Cao, 'Gabor texture representation method for face recognition using the Gamma and generalized Gaussian models', *Image and Vision Computing*, vol. 28, no. 1, pp. 177–187, 2010.

[65] S.-R. Zhou, J.-P. Yin, and J.-M. Zhang, 'Local binary pattern (LBP) and local phase quantization (LBQ) based on Gabor filter for face representation', *Neurocomputing*, vol. 116, pp. 260–264, 2013.

[66] W. Yu *et al.*, 'An improved LBP algorithm for texture and face classification', *Signal, Image and Video Processing*, vol. 8, no. 1, pp. 155–161, 2014.

[67] Z.-H. Huang, W.-J. Li, J. Wang, and T. Zhang, 'Face recognition based on pixel-level and feature-level fusion of the top-level's wavelet sub-bands', *Information Fusion*, vol. 22, pp. 95–104, 2015.

[68] M. Guermoui and M. L. Mekhalfi, 'A Sparse Representation of Complete Local Binary Pattern Histogram for Human Face Recognition', *arXiv preprint arXiv:1605.09584*, 2016.

[69] A. Fathi, P. Alirezazadeh, and F. Abdali-Mohammadi, 'A new Global-Gabor-Zernike feature descriptor and its application to face recognition', *Journal of Visual Communication and Image Representation*, vol. 38, pp. 65–72, 2016.

[70] M. Kan, S. Shan, H. Chang, and X. Chen, 'Stacked progressive auto-encoders (spae) for face recognition across poses', in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1883–1890.

[71] Y. Sun, D. Liang, X. Wang, and X. Tang, 'Deepid3: Face recognition with very deep neural networks', *arXiv preprint arXiv:1502.00873*, 2015.

[72] L. Xinhua and Y. Qian, 'Face recognition based on deep neural network', *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8, no. 10, pp. 29–38, 2015.

[73] Y. Peng, S. Wang, X. Long, and B.-L. Lu, 'Discriminative graph regularized extreme learning machine and its application to face recognition', *Neurocomputing*, vol. 149, pp. 340–353, 2015.

[74] Y. Liu, S. Zhao, Q. Wang, and Q. Gao, 'Learning more distinctive representation by enhanced PCA network', *Neurocomputing*, vol. 275, pp. 924–931, 2018.

[75] I. Masi, A. T. Trần, T. Hassner, J. T. Leksut, and G. Medioni, 'Do we really need to collect millions of faces for effective face recognition?', in *European conference on computer vision*, Springer, 2016, pp. 579–596.

[76] K. Zuiderveld and P. S. Heckbert, 'Graphics gems iv', *San Diego, CA, USA: Academic Press Professional, Inc*, pp. 474–485, 1994.

[77] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, 'The FERET evaluation methodology for face-recognition algorithms', *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[78] T. Choudhury, B. Clarkson, T. Jebara, and A. Pentland, 'Multimodal person recognition using unconstrained audio and video', in *Proceedings, International Conference on Audio-and Video-Based Person Authentication*, Citeseer, 1999, pp. 176–181.

[79] E. Acosta, L. Torres, A. Albiol, and E. Delp, 'An automatic face detection and recognition system for video indexing applications', in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, 2002, p. IV–3644.

[80] B. Moghaddam and M.-H. Yang, 'Learning gender with support faces', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 707–711, 2002.

[81] K. Balci and V. Atalay, 'PCA for gender estimation: Which eigenvectors contribute?', in *2002 International Conference on Pattern Recognition*, IEEE, 2002, pp. 363–366.

[82] Y. Shinohara and N. 1 Otsuf, 'Facial expression recognition using fisher weight maps', in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.*, IEEE, 2004, pp. 499–504.

[83] A. Colmenarez, B. Frey, and T. S. Huang, 'A probabilistic framework for embedded face and facial expression recognition', in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, IEEE, 1999, pp. 592–597.

[84] N. Wang, X. Gao, D. Tao, H. Yang, and X. Li, 'Facial feature point detection: A comprehensive survey', *Neurocomputing*, vol. 275, pp. 50–65, 2018.

[85] Y.-C. Tsai, P.-W. Lai, P.-W. Huang, T.-M. Lin, and B.-F. Wu, 'Vision-based instant measurement system for driver fatigue monitoring', *IEEE Access*, vol. 8, pp. 67342–67353, 2020.

[86] B. H. Prasetio, H. Tamura, and K. Tanno, 'The facial stress recognition based on multi-histogram features and convolutional neural network', in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 2018, pp. 881–887.

[87] V. Azom, A. Adewumi, and J.-R. Tapamo, 'Face and Iris biometrics person identification using hybrid fusion at feature and score-level', in *2015 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*, IEEE, 2015, pp. 207–212.

[88] D. Miao, Z. Sun, and Y. Huang, 'Fusion of multibiometrics based on a new robust linear programming', in *2014 22nd International Conference on Pattern Recognition*, IEEE, 2014, pp. 291–296.

[89] C. Nastar and M. Mitschke, 'Real-time face recognition using feature combination', in *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, IEEE, 1998, pp. 312–317.

[90] S. Phimoltares, C. Lursinsap, and K. Chamnongthai, 'Face detection and facial feature localization without considering the appearance of image context', *Image and Vision Computing*, vol. 25, no. 5, pp. 741–753, 2007.

[91] N. Sharma, S. Saurav, S. Singh, R. Saini, and A. K. Saini, 'A comparative analysis of various image enhancement techniques for facial images', in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2015, pp. 2279–2284.

[92] V. Struc, B. Vesnicer, and N. Pavesic, 'The phase-based gabor fisher classifier and its application to face recognition under varying illumination conditions', in *2008 2nd*

International Conference on Signal Processing and Communication Systems, IEEE, 2008, pp. 1–6.

[93] T. Ojala, M. Pietikäinen, and D. Harwood, 'A comparative study of texture measures with classification based on featured distributions', *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[94] R. Mokni, H. Drira, and M. Kherallah, 'Combining shape analysis and texture pattern for palmprint identification', *Multimedia Tools and Applications*, vol. 76, no. 22, pp. 23981–24008, 2017.

[95] H. Doghmane, H. Bourouba, K. Messaoudi, and A. Bouridane, 'Palmprint recognition based on discriminant multiscale representation', *Journal of Electronic Imaging*, vol. 27, no. 5, p. 053032, 2018.

[96] A. Sharma and K. K. Paliwal, 'Linear discriminant analysis for the small sample size problem: an overview', *International Journal of Machine Learning and Cybernetics*, vol. 6, no. 3, pp. 443–454, 2015.

[97] C. H. Park and H. Park, 'A comparison of generalized linear discriminant analysis algorithms', *Pattern Recognition*, vol. 41, no. 3, pp. 1083–1097, 2008.

[98] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, 'Linear discriminant analysis: A detailed tutorial', *AI communications*, vol. 30, no. 2, pp. 169–190, 2017.

[99] S. Ioffe, 'Probabilistic linear discriminant analysis', in *European Conference on Computer Vision*, Springer, 2006, pp. 531–542.

[100] M. F. Møller, 'A scaled conjugate gradient algorithm for fast supervised learning', *Neural networks*, vol. 6, no. 4, pp. 525–533, 1993.

[101] B. A. Olshausen and D. J. Field, 'Sparse coding with an overcomplete basis set: A strategy employed by V1?', *Vision research*, vol. 37, no. 23, pp. 3311–3325, 1997.

[102] S. Gao, Y. Zhang, K. Jia, J. Lu, and Y. Zhang, 'Single sample face recognition via learning deep supervised autoencoders', *IEEE transactions on information forensics and security*, vol. 10, no. 10, pp. 2108–2118, 2015.

[103] S. Patil, S. Gudasalamani, and N. C. Iyer, 'A survey on Iris recognition system', in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, IEEE, 2016, pp. 2207–2210.

[104] J. G. Daugman, 'High confidence visual recognition of persons by a test of statistical independence', *IEEE transactions on pattern analysis and machine intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.

[105] J. Daugman, 'How iris recognition works', in *The essential guide to image processing*, Elsevier, 2009, pp. 715–739.

[106] S. S. Harakannanavar, C. Prashanth, V. Kanabur, V. I. Puranikmath, and K. Raja, 'An extensive study of issues, challenges and achievements in iris recognition', *Asian Journal of Electrical Sciences*, vol. 8, no. 1, pp. 25–35, 2019.

[107] K. M. A. Alheeti, 'Biometric iris recognition based on hybrid technique', *International Journal on Soft Computing*, vol. 2, no. 4, p. 1, 2011.

[108] M. Y. Shams, M. Z. Rashad, O. Nomir, and R. M. El-Awady, 'Iris recognition based on LBP and combined LVQ classifier', *arXiv preprint arXiv:1111.1562*, 2011.

[109] A. Panganiban, N. Linsangan, and F. Caluyo, 'Wavelet-based feature extraction algorithm for an iris recognition system', *Journal of Information Processing Systems*, vol. 7, no. 3, pp. 425–434, 2011.

[110] R. M. Farouk, 'Iris recognition based on elastic graph matching and Gabor wavelets', *Computer Vision and Image Understanding*, vol. 115, no. 8, pp. 1239–1244, 2011.

[111] G. Sathish, S. Narmadha, S. V. Saravanan, and S. U. Maheswari, 'Multi-algorithmic iris recognition', *International Journal of Computer Applications*, vol. 38, no. 11, pp. 13–21, 2012.

[112] R. Szewczyk, K. Grabowski, M. Napieralska, W. Sankowski, M. Zubert, and A. Napieralski, 'A reliable iris recognition algorithm based on reverse biorthogonal wavelet transform', *Pattern Recognition Letters*, vol. 33, no. 8, pp. 1019–1026, 2012.

[113] S. Zhou and J. Sun, 'A novel approach for code match in iris recognition', in *2013 IEEE/ACIS 12th International Conference on Computer and Information Science (ICIS)*, IEEE, 2013, pp. 123–128.

[114] H. Rai and A. Yadav, 'Iris recognition using combined support vector machine and Hamming distance approach', *Expert systems with applications*, vol. 41, no. 2, pp. 588–593, 2014.

[115] Y. Song, W. Cao, and Z. He, 'Robust iris recognition using sparse error correction model and discriminative dictionary learning', *Neurocomputing*, vol. 137, pp. 198–204, 2014.

[116] Z. Sun, H. Zhang, T. Tan, and J. Wang, 'Iris image classification based on hierarchical visual codebook', *IEEE Transactions on pattern analysis and machine intelligence*, vol. 36, no. 6, pp. 1120–1133, 2013.

[117] N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan, 'Accurate iris segmentation in non-cooperative environments using fully convolutional networks', in *2016 International Conference on Biometrics (ICB)*, IEEE, 2016, pp. 1–8.

[118] H. Proença, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, 'The UBIRIS. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1529–1535, 2009.

[119] N. A. Mashudi and M. J. Nordin, 'A review on iris recognition in non-cooperative environment', in *Proceedings of the 2018 International Conference on Information Science and System*, 2018, pp. 127–132.

[120] U. T. Tania, S. M. A. Motakabber, and M. I. Ibrahimy, 'Edge detection techniques for iris recognition system', in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, 2013, p. 012041.

[121] B. Ammour, T. Bouden, and S. Amira-Biad, 'Multimodal biometric identification system based on the face and iris', in *2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B)*, IEEE, 2017, pp. 1–6.

[122] S. S. Dhage, S. S. Hegde, K. Manikantan, and S. Ramachandran, 'DWT-based feature extraction and radon transform based contrast enhancement for improved iris recognition', *Procedia Computer Science*, vol. 45, pp. 256–265, 2015.

[123] A. F. M. Raffei, H. Asmuni, R. Hassan, and R. M. Othman, 'Feature extraction for different distances of visible reflection iris using multiscale sparse representation of local Radon transform', *Pattern Recognition*, vol. 46, no. 10, pp. 2622–2633, 2013.

[124] P. Viola and M. Jones, 'Rapid object detection using a boosted cascade of simple features', in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, Ieee, 2001, p. I–I.

[125] R. P. Duin and D. M. Tax, 'Experiments with classifier combining rules', in *International Workshop on Multiple Classifier Systems*, Springer, 2000, pp. 16–29.

[126] D. M. Tax, M. Van Breukelen, R. P. Duin, and J. Kittler, 'Combining multiple classifiers by averaging or by multiplying?', *Pattern recognition*, vol. 33, no. 9, pp. 1475–1485, 2000.

[127] M. Indovina, U. Uludag, R. Snelick, A. Mink, and A. Jain, 'Multimodal biometric authentication methods: a COTS approach', in *Proc. of Workshop on Multimodal User Authentication*, Citeseer, 2003, pp. 99–106.

[128] J. Kittler, M. Hatef, R. P. Duin, and J. Matas, 'On combining classifiers', *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

[129] T. K. Ho, J. J. Hull, and S. N. Srihari, 'Decision combination in multiple classifier systems', *IEEE transactions on pattern analysis and machine intelligence*, vol. 16, no. 1, pp. 66–75, 1994.

[130] A. Ross and A. Jain, 'Information fusion in biometrics', *Pattern recognition letters*, vol. 24, no. 13, pp. 2115–2125, 2003.

[131] A. K. Jain and A. Ross, 'Learning user-specific parameters in a multibiometric system', in *Proceedings. International Conference on Image Processing*, IEEE, 2002, p. I–I.

[132] Y. A. Zuev and S. K. Ivanov, 'The voting as a way to increase the decision reliability', *Journal of the Franklin Institute*, vol. 336, no. 2, pp. 361–378, 1999.

[133] L. Lam and C. Y. Suen, 'Optimal combinations of pattern classifiers', *Pattern Recognition Letters*, vol. 16, no. 9, pp. 945–954, 1995.

[134] L. Xu, A. Krzyzak, and C. Y. Suen, 'Methods of combining multiple classifiers and their applications to handwriting recognition', *IEEE transactions on systems, man, and cybernetics*, vol. 22, no. 3, pp. 418–435, 1992.

[135] J. Daugman, 'Combining Multiple Biometrics', The Computer Laboratory, Cambridge University. [Online]. Available: https://www.cl.cam.ac.uk/~jgd1000/combine/

[136] C. Sanderson and B. C. Lovell, 'Multi-region probabilistic histograms for robust and scalable identity inference', in *International conference on biometrics*, Springer, 2009, pp. 199–208.

[137] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, 'Labeled faces in the wild: A database forstudying face recognition in unconstrained environments', in *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*, 2008.

[138] T. Sim, S. Baker, and M. Bsat, ‘The CMU pose, illumination, and expression (PIE) database’, in *Proceedings of fifth IEEE international conference on automatic face gesture recognition*, IEEE, 2002, pp. 53–58.

[139] G.-F. Lu, Z. Jin, and J. Zou, ‘Face recognition using discriminant sparsity neighborhood preserving embedding’, *Knowledge-Based Systems*, vol. 31, pp. 119–127, 2012.

[140] K.-C. Lee, J. Ho, and D. J. Kriegman, ‘Acquiring linear subspaces for face recognition under variable lighting’, *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, no. 5, pp. 684–698, 2005.

Résumé

La multi-biométrie, ou la fusion de plus d'une modalité, d'un échantillon, d'un capteur ou d'un algorithme biométrique, gagne rapidement en popularité en tant que méthode d'amélioration des performances et de la robustesse des systèmes biométriques dans les environnements non contrôlés. Malgré le succès récent de la recherche sur la multi-biométrie, peu de recherches ont été réalisées pour étudier la possibilité de réaliser une fusion multimodale à partir d'un même capteur, en particulier pour la fusion entre le visage et l'iris, qui sont naturellement connexes, surtout après la grande amélioration sur la résolution des caméras les dernières années. Dans ce travail, nous proposons un système : multimodal qui fusionner le visage et l'iris, et multi-algorithme qui fusionner plusieurs méthodes pour l'identification des individus. Cette approche présente de nombreux avantages, notamment la capacité à améliorer les taux de reconnaissance, tout en réduisant le coût d'implémentation et les temps de traitement. Plusieurs études expérimentales sont implémentées pour tester l'efficacité de ce système sur des bases des données courant, les meilleurs taux de reconnaissance étant obtenus en utilisant CLAHE, Gabor, LBP, LPQ, LDA, et l'apprentissage approfondi (l'autoencoder). Les résultats expérimentaux finaux obtenus montre que les performances des deux modalités sont mieux que à modalité unique, et que l'approche proposée donne de meilleurs résultats que les techniques similaires.

Mots-Clés : environnements non contrôlés; reconnaissance de visage ; reconnaissance de l'iris; apprentissage approfondi; Autoencoder; fusion; multimodale; multi algorithme**.**

Abstract

Multi-biometrics, or the fusion of more than one modality, sample, sensor, or biometric algorithm, is rapidly becoming popular as a method to improve the performance and the robustness of biometric systems in uncontrolled environments. Even with the recent success of research on multi-biometrics, little research has been done to study the possibility of achieving multimodal fusion from the same sensor, particularly for the fusion between face and iris, which are naturally related, especially after the great improvement on the resolution of cameras in the last years. In this work, we propose a system multimodal that fuses face and iris, and multi-algorithm that fuses several methods for individual identification. This approach has many advantages, including the ability to improve recognition rates, while reducing implementation cost and processing time. Several experimental studies are implemented to test the effectiveness of this system on current databases, the best recognition rates being obtained using CLAHE, Gabor, LBP, LPQ, LDA, and deep learning (autoencoder). The final experimental results obtained show that the performance of both modalities is better than single modality, and that the proposed approach gives better results than similar techniques.

Keywords: uncontrolled environments; face recognition; iris recognition; deep learning; Autoencoder; fusion; multimodal; multi algorithm.